PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/86, A61K 48/00	A1	(11) International Publication Number: WO 98/22609 (43) International Publication Date: 28 May 1998 (28.05.98)

(21) International Application Number: PCT/US97/21494

(22) International Filing Date: 20 November 1997 (20.11.97)

(30) Priority Data:
08/752,760 20 November 1996 (20.11.96) US

(71) Applicant *(for all designated States except US)*: GENZYME CORPORATION [US/US]; One Mountain Road, Framingham, MA 01701 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: ARMENTANO, Donna, E. [US/US]; 352 Brigthon Street, Belmont, MA 02178 (US). GREGORY, Richard, J. [US/US]; 2 Wintergreen Lane, Westford, MA 01866 (US). SMITH, Alan, E. [GB/US]; 1 Mill Street, Dover, MA 02030 (US).

(74) Agent: SEIDE, Rochelle, K.; Baker & Botts, LLP, 30 Rockefeller Plaza, New York, NY 10112 (US).

(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: CHIMERIC ADENOVIRAL VECTORS

(57) Abstract

A chimeric adenoviral vector is provided that comprises nucleotide sequence of a first adenovirus, wherein all or part of at least one gene of said first adenovirus encoding a protein that facilitates binding of said vector to a target mammalian cell, or internalization thereof within said cell, is replaced by all or part of the corresponding gene from a second adenovirus belonging to subgroup D, said vector further comprising a transgene operably linked to a eucaryotic promoter to allow for expression therefrom in a mammalian cell. Compositions comprising such vectors and methods of using such vectors to deliver transgenes to target mammalian cells, particularly airway epithelial cells, are also provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Description

Chimeric Adenoviral Vectors

Introduction

The present invention relates to chimeric adenoviral vectors, that is, vectors comprising DNA from more than one serotype of adenovirus, which offer enhanced infection efficiency of target cells in order to deliver one or more therapeutically useful nucleotide sequences, including transgenes, therein. Such a nucleotide sequence may comprise a gene not otherwise present in the target cell that codes for a therapeutic and/or biologically active protein, or may represent, for example, an active copy of a gene that is already present in the target cell, but in a defective or deficient form.

Background of the Invention

One of the fundamental challenges now facing medical practicioners is that although the defective genes that are associated with numerous inherited diseases (or that represent disease risk factors including for various cancers) have been isolated and characterized, methods to correct the disease states themselves by providing patients with normal copies of such genes (the technique of gene therapy) are substantially lacking. Accordingly, the development of improved methods of intracellular delivery therefor is of great medical importance. Examples of diseases that it is hoped can be treated by gene therapy include inherited disorders such as cystic fibrosis, Gaucher's disease, Fabry's disease, and muscular dystrophy. Representative of acquired disorders that can be treated are: (1) for cancers: multiple myeloma, leukemias, melanomas, ovarian carcinoma and small cell lung cancer; (2) for cardiovascular conditions: progressive heart failure, restenosis, and hemophilias; and (3) for neurological conditions: traumatic brain injury.

Gene therapy requires successful transfer of nucleic acid to the target cells of a patient. Gene transfer may generally be defined as the process of introducing an expressible polynucleotide (for example a gene, a cDNA, or an mRNA patterned thereon) into a cell. In a particular application of this approach, successful expression of an encoding polynucleotide leads to production in the cells of a normal protein and leads to correction of a disease state associated with an abnormal gene. Therapies based on providing such proteins directly to target cells (protein replacement therapy) have generally proved ineffective since, for example, the cell membrane presents a selectively permeable barrier to entry. Thus there is great interest in alternative methods to cause delivery of therapeutic proteins, especially by transfer of the relevant polynucleotide, often referred to as a transgene.

Viral vectors have been used with increasing frequency to date to deliver transgenes to target cells. Most attempts to use viral vectors for gene therapy have relied on retrovirus-based vectors, chiefly because of their ability to integrate into the cellular genome. However, the disadvantages of retroviral vectors are becoming increasingly clear, including their tropism for dividing cells only, the possibility of insertional mutagenesis upon integration into the cell genome, decreased expression of the transgene over time, rapid inactivation by serum complement, and the possibility of generation of replication-competent retroviruses. See, for example, D. Jolly, et al., Cancer Gene Therapy, 1, 1994, pp. 51-64, and C.P. Hodgson, et al., Bio Technology , 13, 1995, pp. 222-225. Such disadvantages have led to the development of other viral-based vector systems, including those derived from adenoviruses.

Adenovirus (Ad) is a nuclear DNA virus with a genome of about 36 kb, which has been well-characterized through studies in classical genetics and molecular biology. A detailed discussion of adenovirus is found in Thomas Shenk, "Adenoviridae and their Replication", and M. S. Horwitz, "Adenoviruses" , Chapters 67 and 68, respectively, in Virology, B.N. Fields et al., eds., 2nd edition, Raven Press, Ltd., New York, 1996, and reference therein is found to numerous aspects of adenovirus pathology, epidemiology, structure, replication, genetics and classification.

In a simplified form, the adenoviral genome is classified into early (known as E1-E4) and late (known as L1-L5) transcriptional units, referring to the generation of two temporal classes of viral proteins. The demarcation between these events is viral DNA replication.

The human adenoviruses are divided into numerous serotypes (approximately 47, numbered accordingly and classified into 6 subgroups: A, B, C, D, E and F), based upon properties including hemagglutination of red blood cells, oncogenicity, DNA base and protein amino acid compositions and homologies, and antigenic relationships. Additional background information concerning Ad serotype classification, including that for subgroup D, can be found, for example, in F. Deryckere et al., Journal of Virology, 70, 1996, pp. 2832-2841; and A. Bailey et al., Virology, 205, 1994, pp. 438-452, and in other art-recognized references.

Adenoviruses are nonenveloped, regular icosahedrons (having 20 triangular surfaces and 12 vertices) that are about 65-80 nm in diameter. A protein called fiber projects from each of these vertices. The fiber protein is itself generally composed of 3 identical polypeptide chains, although the length thereof varies between serotypes. The protein coat (capsid) is composed of 252 subunits (capsomeres), of which 240 are hexons, and 12 are pentons. Each penton comprises a penton base, on the surface of the capsid, and a fiber protein projecting from the base. The Ad 2 penton base protein, for example, has been determined to be a 8 x 9 nm ring shaped complex composed of 5 identical protein subunits of 571 amino acids each.

Current understanding of adenovirus-cell interactions suggests that adenovirus utilizes two cellular receptors to attach to, and then infect a target cell. It has been further suggested that the fiber protein of an infecting adenovirus first attaches to a receptor, the identity of which is still unknown, and then penton base attaches to a further receptor, often a protein of the alpha integrin family. It has been determined that alpha-integrins often recognize short amino acid sequences on other cellular proteins for attachment pruposes including the tripeptide sequence Arg-Gly-Asp (abbreviated RGD). An RGD sequence is also found in the penton base protein of

adenovirus and is currently understood in the art to mediate attachment of Ad to alpha integrins.

Recombinant adenoviruses have several advantages for use as gene transfer vectors, including tropism for both dividing and non-dividing cells, minimal pathogenic potential, ability to replicate to high titer for preparation of vector stocks, and the potential to carry large inserts (Berkner, K.L.,, Curr. Top. Micro. Immunol. 158:39-66, 1992; Jolly, D., Cancer Gene Therapy 1:51-64, 1994).

The carrying capacity of an adenovirus vector is proportional to the size of the adenovirus genome present in the vector. For example, a capacity of about 8 kb can be created from the deletion of certain regions of the virus genome dispensable for virus growth, e.g., E3, and the deletion of a genomic region such as E1 whose function may be restored in trans from 293 cells (Graham, F.L., J. Gen. Virol. 36:59-72, 1977) or A549 cells (Imler et al., Gene Therapy 3:75-84, 1996). Such E1-deleted vectors are rendered replication-defective, which is desirable for the engineering of adenoviruses for gene transfer. The upper limit of vector DNA capacity for optimal carrying capacity is about 105%-108% of the length of the wild-type genome. Further adenovirus genomic modifications are possible in vector design using cell lines which supply other viral gene products in trans, e.g., complementation of E2a (Zhou et al., J. Virol. 70:7030-7038, 1996), complementation of E4 (Krougliak et al., Hum. Gene Ther. 6:1575-1586, 1995; Wang et al., Gene Ther. 2:775-783, 1995), or complementation of protein IX (Caravokyri et al., J. Virol. 69:6627-6633, 1995; Krougliak et al., Hum. Gene Ther. 6:1575-1586, 1995). Maximal carrying capacity can be achieved using adenoviral vectors deleted for all viral coding sequences (Kochanek et al., Proc. Natl. Acad. Sci. USA 93:5731-5736, 1996; Fisher et al., Virology 217:11-22, 1996).

Transgenes that have been expressed to date by adenoviral vectors include p53 (Wills et al., Human Gene Therapy 5:1079-188, 1994); dystrophin (Vincent et al., Nature Genetics 5:130-134, 1993; erythropoietin (Descamps et al., Human Gene Therapy 5:979-985, 1994; ornithine transcarbamylase (Stratford-Perricaudet et al.,

Human Gene Therapy 1:241-256, 1990; We et al., J. Biol. Chem. 271;3639-3646, 1996;); adenosine deaminase (Mitani et al., Human Gene Therapy 5:941-948, 1994); interleukin-2 (Haddada et al., Human Gene Therapy 4:703-711, 1993); and α1-antitrypsin (Jaffe et al., Nature Genetics 1:372-378, 1992); thrombopoietin (Ohwada et al., Blood 88:778-784, 1996); and cytosine deaminase (Ohwada et al., Hum. Gene Ther. 7:1567-1576, 1996).

The particular tropism of adenoviruses for cells of the respiratory tract has particular relevance to the use of adenovirus in gene therapy for cystic fibrosis (CF), which is the most common autosomal recessive disease in Caucasians. The disease is caused by the presence of one or more mutations in the gene that encodes a protein known as cystic fibrosis transmembrane conductance regulator (CFTR), and which regulates the movement of ions (and therefore fluid) across the cell membrane of epithelial cells, including lung epithelial cells. Abnormal ion transport in airway cells leads to abnormal mucous secretion, inflammmation and infection, tisssue damage, and eventually death. Mutations in the CFTR gene that disturb the cAMP-regulated Cl^- channel in airway epithelia result in pulmonary dysfunction (Zabner et al., Nature Genetics 6:75-83, 1994). Adenovirus vectors engineered to carry the CFTR gene have been developed (Rich et al., Human Gene Therapy 4:461-476, 1993) and studies have shown the ability of these vectors to deliver CFTR to nasal epithelia of CF patients (Zabner et al., Cell 75:207-216, 1993), the airway epithelia of cotton rats and primates (Zabner et al., Nature Genetics 6:75-83, 1994), and the respiratory epithelium of CF patients (Crystal et al., Nature Genetics 8:42-51, 1994). Recent studies have shown that administering an adenoviral vector containing a DNA sequence encoding CFTR to airway epithelial cells of CF patients can restore a functioning chloride ion channel in the treated epithelial cells (Zabner et al., J. Clin. Invest. 97:1504-1511, 1996; U.S. Patent No. 5,670,488 issued September 23, 1997).

Serotype classification is partly based on viral surface protein sequence variation. Because the infectious capabilities of the virus are associated with the surface protein interactions of the virus with cellular proteins, the serotype is an

important determinant of viral entry into target cells, and can account for the infectious heterogeneity of adenovirus serotypes. Most adenoviral vectors have been constructed using adenovirus serotypes from the well-studied group C adenoviruses, especially Ad 2 and Ad 5. However, other adenovirus serotypes display infectious properties that are relevant to the further design of improved adenoviral vectors, for example, those derived from subgroup D, which display enhanced tropism for human airway epithelial cells.

It is widely hoped that gene therapy will provide a long lasting and predictable form of therapy for certain disease states, and it is likely the only form of therapy suitable for many inherited diseases. Although adenoviral vectors are currently in clinical use and have shown therapeutic promise, a need remains to improve the infection efficiency of these vectors in order to further improve their gene transfer capabilities. The present invention addresses this goal.

Summary Of The Invention

The present invention provides for chimeric adenoviral vectors which offer enhanced infection efficiency of target cells for the delivery of one or more transgenes. In a representative aspect of the invention, the vectors comprise nucleotide sequences coding for therapeutically useful proteins and have enhanced tropism for airway epithelial cells.

Accordingly, there are provided chimeric adenoviral vectors comprising nucleotide sequence of a first adenovirus, wherein at least one gene of said first adenovirus encoding a protein that facilitates binding of said vector to a target mammalian cell, or internalization thereof within said cell, is replaced by the corresponding gene from a second adenovirus belonging to subgroup D. These vectors may further comprising a transgene operably linked to a eucaryotic promoter or other regulatory elements to allow for expression therefrom in a mammalian cell. In a representative aspect thereof, the replaced encoding sequence codes for Ad fiber, hexon or penton base.

In a further preferred embodiment of the invention, there are provided chimeric adenoviral vectors comprising nucleotide sequence of a first adenovirus, wherein a portion of a gene thereof encoding a protein that facilitates binding of said vector to a target mammalian cell, or internalization thereof within said cell, is replaced by a portion of the corresponding gene from a second adenovirus belonging to subgroup D. These vectors may further comprising a transgene operably linked to a eucaryotic promoter or other regulatory elements to allow for expression therefrom in a mammalian cell. In a representative aspect thereof, the replaced encoding sequence codes for a portion of Ad fiber, hexon or penton base.

Preferably, the second adenovirus is a member of subgroup D, and the replaced nucleotide sequence encodes a polypeptide selected from the group consisting of Ad fiber, a fragment of Ad fiber, Ad hexon, a fragment of Ad hexon, Ad penton base, and a fragment of Ad penton base. In a preferred embodiment, said second adenovirus is selected from the group consisting of serotypes Ad 9, Ad 15, Ad 17, Ad 19, Ad 20, Ad 22, Ad 26, Ad 27, Ad 28, Ad 30, and Ad 39. In preferred embodiments of the chimeric adenoviral vectors, the first adenovirus is selected from the group consisting of Ad 2, Ad 5, and Ad 12.

The invention is also directed to compositions comprising the chimeric adenoviral vectors of the invention. Additional aspects of the invention include methods to use the chimeric adenoviral vectors of the invention to deliver transgenes to mammalian target cells, for example, to the airway epithelial cells of patients.

A still further representative apsect of the invention involves a method of providing a therapeutic and/or biologically active protein to the airway epithelial cells of a patient by administering to said cells an adenoviral vector comprising elements of an Ad 17 genome, and a transgene encoding said therapeutic protein that is operably linked to a eucaryotic promoter to allow for expression therefrom in a mammalian cell, under conditions whereby the transgene encoding said therapeutic protein is expressed, and therapeutic benefit is produced in said airway epithelial cells.

These and other aspects of the present invention are described in the Detailed Description of the Invention which follows directly.

Brief Description of the Drawings

FIGURE 1 depicts infection of NHBE cells by Ad 2.

FIGURE 2 depicts infection of NHBE cells by Ad 17.

FIGURE 3 plots the result of binding to human nasal polyp epithelial cell isolates by Ad 2 and Ad 17.

FIGURE 4 is a map of the vector Ad2/βgal-2/fiber Ad 17.

FIGURE 5 shows a comparison of the amino acid sequence of penton base from Ad 17 (top) [SEQ ID NO: 4] and Ad 2 (bottom) [SEQ ID NO: 5], and further depicts the variable RGD containing region.

FIGURE 6 depicts an amino acid sequence pileup for penton base from particular Ad serotypes, including f10 (from fowl) [SEQ ID NO: 6 through SEQ ID NO: 10].

FIGURE 7 shows a comparison of the amino acid sequence of fiber from Ad 17 (top) [SEQ ID NO: 11] and Ad 2 (bottom) [SEQ ID NO: 12].

FIGURE 8 depicts an amino acid sequence pileup for fiber from particular Ad serotypes [SEQ ID NO: 11 through SEQ ID NO: 22], including two forms of serotype 40 (40-1 and 40-2) which differ in that one variant has two (but non-identical) copies of the fiber gene.

FIGURE 9 shows the infection efficiency of colon cancer cell lines by adenovirus serotypes.

FIGURE 10 shows the infection efficiency of cancer cell lines by adenovirus serotypes.

Provided in the Sequence Listing attached hereto are also:

SEQ ID NO: 1, the complete nucleotide sequence of Ad 17;

SEQ ID NO: 2, the complete encoding nucleotide sequence for Ad 17 fiber;

SEQ ID NO: 3, the complete encoding nucleotide sequence for Ad 17 penton base.

Detailed Description of the Invention

The present invention provides for chimeric adenoviral vectors comprising nucleotide sequence of a first adenovirus, wherein at least one gene of said first adenovirus encoding a protein that facilitates binding of said vector to a target mammalian cell, or internalization thereof within said cell, is replaced by the corresponding gene from a second adenovirus belonging to subgroup D, said vectors further comprising a transgene operably linked to a eucaryotic promoter to allow for expression therefrom in a mammalian cell. In a representative aspect thereof, the replaced encoding sequence correspond to the gene encoding the Ad fiber, hexon or penton base proteins, or combinations thereof.

In a further preferred embodiment of the invention, there are provided chimeric adenoviral vectors comprising nucleotide sequence of a first adenovirus, wherein a portion of a gene thereof encoding a protein that facilitates binding of said vector to a target mammalian cell, or internalization thereof within said cell, is replaced by a portion of the corresponding gene from a second adenovirus belonging to subgroup D, said vectors further comprising a transgene operably linked to a eucaryotic promoter to allow for expression therefrom in a mammalian cell. In a representative aspect thereof, the replaced encoding sequence codes for a portion of the Ad fiber, hexon or penton base proteins, or combinations thereof. Where a portion of a gene from a second adenovirus is used to construct a chimeric adenoviral vector, such sequence will have a length sufficient to confer a desired serotypic-specific virus-cell interaction to the vector.

The present invention involves the recognition that adenoviral vectors that are either based substantially upon the genome of Ad serotypes classified in subgroup D, or that contain certain Ad-protein encoding polynucleotide sequences of subgroup D adenovirus, are particularly effective at binding to, and internalizing within, human

cells, such that therapeutic transgenes included in the adenoviral vector are efficiently expressed. This discovery is particularly surprising given that adenovirus serotypes of subgroup D are not clinically associated with human respiratory disease, and that, for example association with conjunctivitis is more typical. The recognition of this tropism is of particular relevance for the treatment by gene therapy of recognized disease states such as cystic fibrosis or α 1-antitrypsin deficiency.This discovery is particularly surprising given that adenovirus serotypes of subgroup D are not clinically associated with human respiratory disease, and that, for example association with conjunctivitis is more typical. The recognition of this tropism is of particular relevance for the treatment by gene therapy of recognized disease states such as cystic fibrosis or α 1-antitrypsin deficiency.

In a representative aspect of the invention, the adenoviral vectors further comprise nucleotide sequences coding for one or more transgenes and have enhanced tropism for airway epithelial cells. Preferably, the chimeric adenoviral vectors are replication-defective, a feature which contributes to the enhanced safety of adenoviral vectors administered to individuals.

Preferably, the second adenovirus is a member of subgroup D, and the replaced nucleotide sequence encodes a polypeptide selected from the group consisting of Ad fiber, a fragment of Ad fiber, Ad hexon, a fragment of Ad hexon, Ad penton base, and a fragment of Ad penton base. In a preferred embodiment, said second adenovirus is selected from the group consisting of serotypes Ad 9, Ad 15, Ad 17, Ad 19, Ad 20, Ad 22, Ad 26, Ad 27, Ad 28, Ad 30, and Ad 39. In a most preferred embodiment, the second adenovirus is Ad 17. In other preferred embodiments of the chimeric adenoviral vectors, the first adenovirus is selected from the group consisting of Ad 2, Ad 5, and Ad 12.

There is substantial evidence that any reported transforming properties of the E4 region of certain subgroup D serotypes do not extend to Ad serotypes whose use is preferred according to the practice of the present invention (see, for example, R. Javier

et al., Science, 257, 1992, pp. 1267-1271). It is expected also that, for example, individual ORFs of subgroup D E4 region, such as ORF1, could be deleted.

Additional aspects of the invention include methods to provide biologically active and/or therapeutic proteins to mammalian cells, including, but not limited to, the airway epithelial cells of individuals, in order to provide phenotypic benefit. According to this aspect of the invention, chimeric adenoviral vectors are used in which a nucleotide sequence of a first adenovirus is replaced by the corresponding nucleotide sequence of a second adenovirus. Preferably, the second adenovirus is a member of subgroup D, and the replaced nucleotide sequence encodes a polypeptide encoding all or part of Ad fiber, Ad hexon, or Ad penton base, or combinations thereof.

A still further representative aspect of the invention involves providing a biologically active and/or therapeutic protein in the airway epithelial cells of a patient by administering to said cells an adenoviral vector comprising elements of an Ad 17 genome, and a transgene encoding said protein that is operably linked to a eucaryotic promoter to allow for expression therefrom in a mammalian cell, under conditions whereby the transgene encoding said protein is expressed, and the desired phenotypic benefit is produced in said airway epithelial cells. According to the practice of the invention, it is preferred that an chimeric adenovirus vector utilized to deliver a transgene to the respiratory epithelium (including that of the nasal airway, trachea, and bronchi and alveoli of the lung), or to other tissues of the body, comprise serotypes within subgroup D, as such classification is recognized in the art.

In order to construct the chimeric adenoviral vectors of the invention, reference may be made to the substantial body of literature on how such vectors may be designed, constructed and propagated using techniques from molecular biology and microbiology that are well-known to the skilled artisan. Specific examples of adenoviral vector genomes which can be used as the backbone for a chimeric adenoviral vector of the invention include, for example, Ad2/CFTR-1 and Ad2/CFTR-2 and others described in U. S. Patent No. 5,670,488, issued September 23, 1997

(incorporated herein by reference). Such vectors may include deletion of the E1 region, partial or complete deletion of the E4 region, and deletions within, for example, the E2 and E3 regions. Within the scope of the invention are, for example, chimeric vectors which contain an Ad 2 backbone with one or more Ad 17 capsid proteins or fragments thereof in the virus. Other adenoviral vector genomic designs which can be used in the chimeric adenoviral vectors of the invention include those derived from allowed U.S. Patent Application Serial No. 08/409,874, filed March 24, 1995, and allowed U.S. Patent Application Serial No. 08/540,077, filed October 6, 1995 (both incorporated herein by reference).

To construct the recombinant chimeric adenoviral vectors of the invention which contain a transcription unit, the skilled artisan can use the standard techniques of molecular biology to engineer a transgene or a capsid protein into a backbone vector genome (Berkner, K.L., Curr. Top. Micro. Immunol. 158:39-66, 1992). For example, a plasmid containing a transgene and any operably linked regulatory elements inserted into an adenovirus genomic fragment can be co-transfected with a linearized viral genome derived from an adenoviral vector of interest into a recipient cell under conditions whereby homologous recombination occurs between the genomic fragment and the virus. Preferably, a transgene is engineered into the site of an E1 deletion. As a result, the transgene is inserted into the adenoviral genome at the site in which it was cloned into the plasmid, creating a recombinant adenoviral vector. The chimeric adenoviral vectors can also be constructed using standard ligation techniques, for example, removing a restriction fragment containing a fiber gene from a first adenovirus and ligating into that site a restriction fragment containing a fiber gene from a second adenovirus. A representative example of a chimeric adenoviral vector of the invention is Ad2/βgal-2 fiber 17 (exemplified in Example 6).

Construction of the chimeric adenoviral vectors can be based on adenovirus DNA sequence information widely available in the field, e.g., nucleic acid sequence databases such as GenBank.

Preparation of replication-defective chimeric adenoviral vector stocks can be accomplished using cell lines that complement viral genes deleted from the vector, e.g., 293 or A549 cells containing the deleted adenovirus E1 genomic sequences. The use of HER3 cells (human embryonic retinoblasts transformed by Ad 12), as a complementing cell line is of note. After amplification of plaques in suitable complementing cell lines, the viruses can be recovered by freeze-thawing and subsequently purified using cesium chloride centrifugation. Alternatively, virus purification can be performed using chromatographic techniques, e.g., as set forth in International Application No. PCT/US96/13872, filed August 30, 1996, incorporated herein by reference.

Titers of replication-defective chimeric adenoviral vector stocks can be determined by plaque formation in a complementing cell line, e.g., 293 cells. End-point dilution using an antibody to the adenoviral hexon protein may be used to quantitate virus production or infection efficiency of target cells (Armentano et al., Hum. Gene Ther. 6:1343-1353, 1995, incorporated herein by reference).

Transgenes which can be delivered and expressed from a chimeric adenoviral vector of the invention include, but are not limited to, those encoding enzymes, blood derivatives, hormones, lymphokines such as the interleukins and interferons, coagulants, growth factors, neurotransmitters, tumor suppressors, apoliproteins, antigens, and antibodies, and other biologically active proteins. Specific transgenes which may be encoded by the chimeric adenoviral vectors of the invention include, but are not limited to, cystic fibrosis transmembrane regulator (CFTR), dystrophin, glucocerebrosidase, tumor necrosis factor, p53, p21, herpes simplex thymidine kinase and gancyclovir, retinoblastoma (Rb), and adenosine deaminase (ADA). Transgenes encoding antisense molecules or ribozymes are also within the scope of the invention. The vectors may contain one or more transgenes under the control of one or more regulatory elements.

In addition to containing the DNA sequences encoding one or more transgenes, the chimeric adenoviral vectors of the invention may contain any

expression control sequences such as a promoter or enhancer, a polyadenylation element, and any other regulatory elements that may be used to modulate or increase expression, all of which are operably linked in order to allow expression of the transgene. The use of any expression control sequences, or regulatory elements, which facilitate expression of the transgene is within the scope of the invention. Such sequences or elements may be capable of generating tissue-specific expression or be susceptible to induction by exogenous agents or stimuli.

Infection of target cell by the chimeric adenoviral vectors of the invention may also be facilitated by the use of cationic molecules, such as cationic lipids as disclosed in PCT Publication No. WO96/18372, published June 20, 1996, incorporated herein by reference.

Cationic amphiphiles have a chemical structure which encompasses both polar and non-polar domains so that thc molecule can simultaneously facilitate entry across a lipid membrane with its non-polar domain while its cationic polar domain attaches to a biologically useful molecule to be transported across the membrane.

Cationic amphiphiles which may be used to form complexes with the chimeric adenoviral vectors of the invention include, but are not limited to, cationic lipids, such as DOTMA (Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417, 1987) (N-[1-(2,3-dioletloxy)propyl]-N,N,N - trimethylammonium chloride); DOGS (dioctadecylamidoglycylspermine) (Behr et al., Proc. Natl. Acad. Sci. USA 86:6982-6986, 1989); DMRIE (1,2-dimyristyloxypropyl-3-dimethyl-hydroxyethyl ammonium bromide) (Felgner et al., J. Biol. Chem. 269:2550-2561, 1994; and DC-chol (3B [N-N', N'-dimethylaminoethane) -carbamoyl] cholesterol) (U.S. Patent No. 5, 283,185 to Epand et al.). The use of other cationic amphiphiles recognized in the art or which come to be discovered is within the scope of the invention.

In preferred embodiments of the invention, the cationic amphiphiles useful to complex with and facilitate transfer of the vectors of the invention are those lipids which are described in PCT Publication No. WO96/18372, published June 20, 1996, which is incorporated herein by reference. Preferred cationic amphiphiles described

herein to be used in the delivery of the plasmids and/or viruses are GL-53, GL-67, GL-75, GL-87, GL-89, and GL-120, including protonated, partially protonated, and deprotonated forms thereof. Further embodiments include the use of non-T-shaped amphiphiles as described on pp. 22-23 of the aforementioned PCT application, including protonated, partially protonated and deprotonated forms thereof. Most preferably, the cationic amphiphile which can be used to deliver the vectors of the invention is spermine cholesterol carbamate (GL-67).

In the formulation of compositions comprising the chimeric adenoviral vectors of the invention, one or more cationic amphiphiles may be formulated with neutral co-lipids such as dileoylphosphatidylethanolamine (DOPE) to facilitate delivery of the vectors into a cell. Other co-lipids which may be used in these complexes include, but are not limited to, diphytanoylphosphatidylethanolamine, lyso-phosphatidylethanolamines, other phosphatidylethanolamines, phosphatidylcholines, lyso-phosphatidylcholines and cholesterol. A preferred molar ratio of cationic amphiphile to colipid is 1:1. However, it is within the scope of the invention to vary this ratio, including also over a considerable range. In a preferred embodiment of the invention, the cationic amphiphile GL-67 and the neutral co-lipid DOPE are combined in a 1:2 molar ratio, respectively, before complexing with a chimeric adenoviral vector for delivery to a cell.

In the formulation of complexes containing a cationic amphiphile with a chimeric adenoviral vector, a preferred range of 10^7 - 10^{10} infectious units of virus may be combined with a range of 10^4 - 10^6 cationic amphiphile molecules/viral particle.

The infection efficiency of the chimeric adenoviral vectors of the invention may be assayed by standard techniques to determine the infection of target cells. Such methods include, but are not limited to, plaque formation, end-point dilution using, for example, an antibody to the adenoviral hexon protein, and cell binding assays using radiolabelled virus. Improved infection efficiency may be characterized as an increase in infection of at least an order of magnitude with reference to a control virus. Where

a chimeric adenoviral vector encodes a marker or other transgene, relevant molecular assays to determine expression include the measurement of transgene mRNA, by, for example, Northern blot, S1 analysis or reverse transcription-polymerase chain reaction (RT-PCR). The presence of a protein encoded by a transgene may be detected by Western blot, immunoprecipitation, immunocytochemistry, or other techniques known to those skilled in the art. Marker-specific assays can also be used, such as X-gal staining of cells infected with a chimeric adenoviral vector encoding β-galactosidase.

In order to determine transgene expression and infection efficiency in vivo using the constructs and compositions of the invention, animal models may be particularly relevant in order to assess transgene persistence against a background of potential host immune response. Such a model may be chosen with reference to such parameters as ease of delivery, identity of transgene, relevant molecular assays, and assessment of clinical status. Where the transgene encodes a protein whose lack is associated with a particular disease state, an animal model which is representative of the disease state may optimally be used in order to assess a specific phenotypic result and clinical improvement. However, it is also possible that particular chimeric adenoviral vectors of the invention display enhanced infection efficiency only in human model systems, e.g., using primary cell cultures, tissue explants, or permanent cell lines. In such circumstances where there is no animal model system available in which to model the infection efficiency of a chimeric adenoviral vector with respect to human cells, reference to art-recognized human cell culture models will be most relevant and definitive.

Relevant animals in which the chimeric adenoviral vectors may be assayed include, but are not limited to, mice, rats, monkeys, and rabbits. Suitable mouse strains in which the vectors may be tested include, but are not limited to, C3H, C57Bl/6 (wild-type and nude) and Balb/c (available from Taconic Farms, Germantown, New York).

Where it is desirable to assess the host immune response to vector administration, testing in immune-competent and immune-deficient animals may be

compared in order to define specific adverse responses generated by the immune system. The use of immune-deficient animals, e.g., nude mice, may be used to characterize vector performance and persistence of transgene expression, independent of an acquired host response.

In a particular embodiment where the transgene is the gene encoding cystic fibrosis transmembrane regulator protein (CFTR) which is administered to the respiratory epithelium of test animals, expression of CFTR may be assayed in the lungs of relevant animal models, for example, C57Bl/6 or Balb/c mice, cotton rats, or Rhesus monkeys. Molecular markers which may used to determine expression include the measurement of CFTR mRNA, by, for example, Northern blot, S1 analysis or RT-PCR. The presence of the CFTR protein may be detected by Western blot, immunoprecipitation, immunocytochemistry, or other techniques known to those skilled in the art. Such assays may also be used in tissue culture where cells deficient in a functional CFTR protein and into which the chimeric adenoviral vectors have been introduced may be assessed to determine the presence of functional chloride ion channels - indicative of the presence of a functional CFTR molecule.

The chimeric adenoviral vectors of the invention have a number of in vivo and in vitro utilities. The vectors can be used to transfer a normal copy of a transgene encoding a biologically active protein to target cells in order to remedy a deficient or dysfunctional protein. The vectors can be used to transfer marked transgenes (e.g., containing nucleotide alterations) which allow for distinguishing expression levels of a transduced gene from the levels of an endogenous gene. The chimeric adenoviral vectors can also be used to define the mechanism of specific viral protein-cellular protein interactions that are mediated by specific virus surface protein sequences. The vectors can also be used to optimize infection efficiency of specific target cells by adenoviral vectors, for example, using a chimeric adenoviral vector containing Ad 17 fiber protein to infect human nasal polyp cells. Where it is desirable to use an adenoviral vector for gene transfer to cancer cells in an individual, a chimeric adenoviral vector can be chosen which selectively infects the specific type of target

cancer cell and avoids promiscuous infection. Where primary cells are isolated from a tumor in an individual requiring gene transfer, the cells may be tested against a panel of chimeric adenoviral vectors to select a vector with optimal infection efficiency for gene delivery. The vectors can further be used to transfer tumor antigens to dendritic cells which can then be delivered to an individual to elicit an anti-tumor immune response. Chimeric adenoviral vectors can also be used to evade undesirable immune responses to particular adenovirus serotypes which compromise the gene transfer capability of adenoviral vectors.

The present invention is further directed to compositions containing the chimeric adenoviral vectors of the invention which can be administered in an amount effective to deliver one or more desired transgenes to the cells of an individual in need of such molecules and cause expression of a transgene encoding a biologically active protein to achieve a specific phenotypic result. The cationic amphiphile-plasmid complexes or cationic amphiphile-virus complexes may be formulated into compositions for administration to an individual in need of the delivery of the transgenes.

The compositions can include physiologically acceptable carriers, including any relevant solvents. As used herein, "physiologically acceptable carrier" includes any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like. Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the compositions is contemplated.

Routes of administration for the compositions containing the chimeric adenoviral vectors of the invention include conventional and physiologically acceptable routes such as direct delivery to a target organ or tissue, intranasal, intravenous, intramuscular, subcutaneous, intradermal, oral and other parenteral routes of administration.

The invention is further directed to methods for using the compositions of the invention in vivo or ex vivo applications in which it is desirable to deliver one or more

transgenes into cells such that the transgene produces a biologically active protein for a normal biological or phenotypic effect. In vivo applications involve the direct administration of one ore more chimeric adenoviral vectors formulated into a composition to the cells of an individual. Ex vivo applications involve the transfer of a composition containing the chimeric adenoviral vectors directly to autologous cells which are maintained in vitro, followed by readministration of the transduced cells to a recipient.

Dosage of the chimeric adenoviral vector to be administered to an individual for expression of a transgene encoding a biologically active protein and to achieve a specific phenotypic result is determined with reference to various parameters, including the condition to be treated, the age, weight and clinical status of the individual, and the particular molecular defect requiring the provision of a biologically active protein. The dosage is preferably chosen so that administration causes a specific phenotypic result, as measured by molecular assays or clinical markers. For example, determination of the infection efficiency of a chimeric adenoviral vector containing the CFTR transgene which is administered to an individual can be performed by molecular assays including the measurement of CFTR mRNA, by, for example, Northern blot, S1 or RT-PCR analysis or the measurement of the CFTR protein as detected by Western blot, immunoprecipitation, immunocytochemistry, or other techniques known to those skilled in the art. Relevant clinical studies which could be used to assess phenotypic results from delivery of the CFTR transgene include PFT assessment of lung function and radiological evaluation of the lung. Demonstration of the delivery of a transgene encoding CFTR can also be demonstrated by detecting the presence of a functional chloride channel in cells of an individual with cystic fibrosis to whom the vector containing the transgene has been administered (Zabner et al., J. Clin. Invest. 97:1504-1511, 1996). Transgene expression in other disease states can be assayed analogously, using the specific clinical parameters most relevant to the condition.

Dosages of a chimeric adenoviral vector which are effective to provide expression of a transgene encoding a biologically active protein and achieve a specific phenotypic result range from approximately 10^8 infectious units (I.U.) to 10^{11} I.U. for humans.

It is especially advantageous to formulate parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subjects to be treated, each unit containing a predetermined quantity of active ingredient calculated to produce the specific phenotypic effect in association with the required physiologically acceptable carrier. The specification for the novel dosage unit forms of the invention are dictated by and directly depend on the unique characteristics of the chimeric adenoviral vector and the limitations inherent in the art of compounding. The principal active ingredient (the chimeric adenoviral vector) is compounded for convenient and effective administration in effective amounts with the physiologically acceptable carrier in dosage unit form as discussed above.

Maximum benefit and achievement of a specific phenotypic result from administration of the chimeric adenoviral vectors of the invention may require repeated administration. Such repeated administration may involve the use of the same chimeric adenoviral vector, or, alternatively, may involve the use of different chimeric adenoviral vectors which are rotated in order to alter viral antigen expression and decrease host immune response.

The practice of the invention employs, unless otherwise indicated, conventional techniques of protein chemistry, molecular virology, microbiology, recombinant DNA technology, and pharmacology, which are within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Current Protocols in Molecular Biology, Ausubel et al., eds., John Wiley & Sons, Inc., New York, 1995, and Remington's Pharmaceutical Sciences, 17th ed., Mack Publishing Co., Easton, PA, 1985.

The invention is further illustrated by the following specific examples which are not intended in any way to limit the scope of the invention.

Examples

Example 1 Infection of NHBE cells by adenovirus serotypes of subgroup D

Normal human bronchial epithelial ("NHBE") cells were obtained from Clonetics (San Diego, CA), and plated on Costar (Cambridge, MA) Transwell-Clear polyester membranes that were pre-coated with human placental collagen. The wells were placed in a cluster plate and cells were fed every day for one week by changing the medium in both the well and the plate. After one week the media was removed from the wells to create an air-liquid interface, and the cells were then fed only by changing the medium in the cluster plate, every other day for one week. Cells were infected at an moi of 1 by adding virus (see below) to the transwell, followed by an incubation time of 1.5-2 hours. At the end of the incubation period, the medium was removed and the cells were gently rinsed with fresh medium. Thirty-six hours post-infection the cells were fixed with 1:1 acetone:methanol, permeablized with a solution of 0.05% Tween 20 in PBS, and stained with FITC labeled anti-hexon antibody (Chemicon, Temecula, CA) to visualize cells that had been productively infected (i.e. to visualize virus replication). Cells were also subjected to the DAPI staining procedure in order to visualize the total number of nuclei. The results could be readily determined upon simple inspection.

Wild type Ad serotypes within subgroup D that were tested included 9, 15, 17, 19, 20, 22, 26, 27, 28, 30, and 39 (all from the American Type Culture Collelction, Rockville, MD). An Ad 2 (obtained as DNA from BRL, Gaithersburg, MD, and used to transfect 293 cells in order to generate virus stock) was used as a control. Infection observed with all of the subgroup D serotypes was superior to that observed with Ad 2, with the best results being achieved with Ad 9, Ad 17, Ad 20, Ad 22, and Ad 30.

Additionally, it was determined that each of the above-mentioned serotypes of subgroup D was more effective in the NHBE cell assay under similar circumstances than any other serotype tested than belongs to a subgroup other than D. In this regard, the following serotypes were also tested: 31(subgroup A); 3(subgroup B); 7(subgroup B); 7a(subgroup B); 14(subgroup B); 4(subgroup E); and 41(subgroup F). In a further experiment, serotype 35 (subgroup A) may have performed as well as the least effective members of subgroup D that were tested.

Example 2 Infection of clinical isolate bronchial epithelial cells

Following generally the procedures of Example 1, human bronchial epithelial cells recovered from healthy human volunteers were infected with either Ad 2 (as above, Ad 2 DNA was obtained from BRL, and this DNA was used to transfect 293 cells to generate virus) (Figure 1), or Ad 17 (from ATCC) (Figure 2), all at an moi of 50. Cells were left in contact with virus for 30 minutes, 3 hours, or 12 hours.

The increased tropism of Ad 17 for human bronchial epithelial cells, compared with Ad 2, is readily apparent upon inspection of Figures 1 and 2. In the Figures, the right hand columns (panels D, E, and F, stained in blue) show total numbers of cells present (from DAPI staining as above), whereas the left hand columns (panels A, B, and C, stained in green) quantify adenovirus hexon protein present in the infected cells (from FITC-labeled anti-hexon anitbody, as above). Panels A and D result from 30 minute incubation times, panels B and E result from 3 hour incubation times, and panels C and F result from 12 hour incubation times. As measured by the technique employed, infection of airway epithelia by Ad 17 is at least 50 fold greater than by Ad 2 for the thirty minute incubation time.

Example 3 Binding of Ad 2 and Ad 17 to human nasal polyp cell isolates

293 cells, a complementing cell line developed by Graham et al. (see Gen. Virol. , 36, 1977, pp. 59-72), were infected with either wild type Ad 2 or wild type Ad 17. Five hours post-infection the media was removed and replaced with methionine

free media containing S^{35} metabolic label (Amersham). After an additional six hours, fresh media was added and the labeling was allowed to proceed for a total of 18 hours, after which the S^{35} media was removed and replaced with fresh media. Thirty hours post-infection the cells were harvested and lysed and the labeled Ad 2 or Ad 17 viruses were purified by CsCl gradient centrifugation. The recovered viruses were then used in an assay to determine their relative binding efficiency on human nasal polyp cells.

In order to perform the assay, ciliated human airway epitehlial cells were recovered from nasal polyps of healthy volunteers. The results from two such isolates, NP-14 and NP-15, are reported here (see Figure 3). Radiolabeled virus was then incubated with the isolated cells in wells for specified times (5 or 30 minutes, see Figure 3). The cells were then rinsed and measured for radioactivity. Binding as reported in Figure 3 indicates the percent of input radioactivity that is cell associated. It was determined that for both cell isolate populations, using either 5 or 30 minute incubations, cell associated radioactivity was 10-fold enhanced if Ad 17 rather than Ad 2 was used.

Example 4 Fiber competition

A549 cells (a human lung carcinoma line, obtained from the American Type Culture Collection as ATCC CCL-185) were plated at 3 x 10^4 cells per well in 96-well dishes. Since the number of receptor sites for adenovirus fiber on the cell surface has been estimated to be approximately 10^5 receptors per cell, the receptors in the plated cells were saturated, in this example, with 0.1 μg of purified full length Ad 2 fiber protein (obtained from Paul Freimuth, Brookhaven National Laboratory, Upton, NY), which corresponds to approximately 100 molecules of fiber per receptor. Cells were incubated with Ad 2 fiber in PBS for two hours at 37°C.

The cells were subsequently infected at an moi of 1 (using either Ad 2 provided as above, or wild type Ad 17) for one hour, after which the cells were rinsed, and fresh mediium was added. Control cultures were incubated with PBS with no added protein for two hours and then subsequently infected as described above. Forty hours post-infection the cells were fixed with 1:1 acetone:methanol, permeablized with 0.05% Tween 20 in PBS and stained with FITC labeled anti- Ad 2 hexon antibody, as described in Example 1. As determined by this assay, the number of cells infected (stained) with Ad 2 was reduced by approximately 90% in cultures that were pre-incubated with Ad 2 fiber as compared to control cultures. However, no effect on Ad 17 infection was observed by the pre-incubation of A549 cells with full length Ad 2 fiber.

Example 5 Use of Ad 2 fiber knob in a binding competition experiment with Ad 2

Further competition experiments were performed with Ad 2 and Ad 17 fiber knobs that had been expressed and purified from E. coli. DNA sequences encoding both protein fragments were designed so that the fiber knobs expressed therefrom would contain histidine tags in order to permit nickel- column purification. The yield of soluble fiber knob trimer, purified by the Ni-NTA method (Qiagen, Chatsworth, CA) , was ~25μg/50ml culture. A significant portion of the total knob protein expressed appeared to remain in a monomeric (and insoluble) form. The soluble trimeric material obtained was used for a preliminary competition experiment. Wild type Ad 2 and Ad 17 were used to infect A549 cells, or cells that had been pre-incubated with excess (about 100 molecules of trimer per receptor) Ad 2 fiber knob or Ad 17 fiber knob. The results indicated that Ad 2 fiber knob, but not Ad 17 knob, could block Ad 2 infection. Additionally, Ad 17 infection was not blocked by E. coli-expressed fiber knobs of either serotype, suggesting that the mechanism of Ad 2 and Ad 17 infections is different.

Example 6 Construction of the chimeric vector Ad2/βgal-2/fiber Ad 17

The vector Ad2/βgal-2 was constructed as follows. A CMV§gal expression cassette was constructed in a pBR322-based plasmid that contained Ad 2 nucleotides 1-10,680 from which nucleotides 357-3328 were deleted. The deleted sequences were replaced with (reading from 5' to 3'): a cytomegalovirus immediate early promoter (obtained from pRC/CMV, Invitrogen), lacZ gene encoding §-galactosidase with a nuclear localization signal, and an SV40 polyadenylation signal (nucleotides 2533-2729). The resulting plasmid was used to generate Ad2/βgal-2 by recombination with Ad2E4ORF6 (D. Armentano et al., Human Gene Therapy , 6, 1995, pp 1343 -1353).

A chimeric Ad2/βgal-2/fiber Ad 17 viral vector (Figure 4) was then contructed as follows. pAdORF6 (D. Armentano et al., Human Gene Therapy , 6, 1995, pp 1343 -1353 was cut with Nde and BamHI to remove Ad 2 fiber coding and polyadenylation signal sequences (nucleotides 20624-32815). An NdeI-BamHI fragment containing Ad 17 fiber coding sequence (nucleotides 30984-32095) was generated by PCR and ligated along with an SV40 polyadenylation signal into NdeI-BamHI cut pAdORF6 to generate pAdORF6fiber17. This plasmid was cut with PacI and then ligated to PacI-cut Ad2/βgal-2 DNA to generate Ad2/βgal-2fiber 17. Any desired transgene may be substituted in this construct for the reporter gene.

A similar construct can be prepared using a DNA sequence that encodes Ad 17 penton base instead of Ad 17 fiber. Alternatively, only a subregion of the penton base of Ad 2 need be subject to replacement, such as by inserting into the vector a nucleotide encoding sequence corresponding to any amino acid subsequence of Ad 17 penton base amino acids 283-348 (see the marked sequence in Figure 5A) in replacement for any subsequence of Ad 2 penton base amino acids 290-403. Preferrably, the replaced sequence of Ad 2 and the inserted sequence of Ad 17 includes the RGD domain of each. Use of nucleotide sequence corresponding to penton base amino acid sequence for other subgroup D serotypes is also within the

practice of the invention. It is also within the scope of the invention to replace a subregion of the fiber protein in the Ad 2 vector with a subregion from another adenovirus serotype, for example, Ad 17.

Example 7 Ad2/βgal-2f17 shows increased infection efficiency on human airway explants

Both human and monkey trachea explants, about 1 cm^2, were placed on top of an agar support. Each explant was infected at an moi of 200 of either Ad2/βgal-2 or Ad2/βgal-2f17 assuming a cell density of 1×10^6 per cm^2 of explant. Explants were exposed to virus for three hours and were then rinsed with NHBE media. Two days post-infection explants were stained with X-gal and infection efficiency was assessed. On the monkey explants Ad2/βgal-2 gave rise to a higher infection efficiency than Ad2/βgal-2f17. Patches of stained cells were detected in explants exposed to Ad2/βgal-2 but very few cells stained in explants exposed to Ad2/βgal-2f17. A different result was obtained on human trachea explants. On these explants Ad2/βgal-2f17 infection gave rise to a much higher infection efficiency than Ad2/βgal-2 infection. Approximately 5-10% of the cells in explants exposed to Ad2/βgal-2f17 stained with X-gal whereas very few cells were stained in explants exposed to Ad2/βgal-2. No background staining was observed in either monkey or human explants that were not exposed to virus.

The results indicate that the exchange of Ad 2 fiber for Ad 17 fiber in Ad2/βgal-2f17 was suffficient to significantly increase infection efficiency of human tracheal airway cells by an adenovirus type 2 based vector.

Example 8 Adenovirus subgroup screening on human cancer cell lines

Identification of adenovirus subgroup that best infects a particular tumor type may be useful in designing vectors to optimally target cancer cells in vivo. In order to determine the adenovirus subgroup that best infects a particular type of cancer cell, cancer cells were seeded into a 96 well plate and infected with and moi of 5. Infection

efficiency was determined by staining of infected cells using an anti-hexon antibody. The adenovirus subgroups were represented by the following serotypes: A: Ad 31; B: Ad 3; C: Ad 2; D: Ad 17; E: Ad 4; and F: Ad 41.

Subgroup D (Ad 17) has a significantly higher infection rate of the colon cancer cell line CaCo-2 than other cell types, with an infection rate of 70%, while Ad 2 only infected 20% of the cells (Figure 9).

Subgroup D (Ad 17) was effective in infecting ovarian cancer cell line SK-OV3. Infection was measured at 90% (FIgure 10).

Sequence Listing

Included herewith on the following pages are informal copies of SEQ ID NO: 1 through SEQ ID NO: 3.

```
   1 CATCATCAAT AATATACCCC ACAAAGTAAA CAAAAGTTAA TATGCAAATG AGGTTTTAAA
  61 TTTAGGGCGG GGCTACTGCT GATTGGCCGA GAAACGTTGA TGCAAATGAC GTCACGACGC
 121 ACGGCTAACG GTCGCCGCGG AGGCGTGGCC TAGCCCGGAA GCAAGTCGCG GGGCTGATGA
 181 CGTATAAAAA AGCGGACTTT AAACCCGGAA ACGGCCGATT TTCCCGCGGC CACGCCCGGA
 241 TATGAGGTAA TTCTGGGCGG ATGCAAGTGA AATTAGGTCA TTTTGGCGCG AAAACTGAAT
 301 GAGGAAGTGA AAAGTGAAAA ATACCGGTCC CGCCCAGGGC GGAATATTTA CCGAGGGCCG
 361 AGAGACTTTG ACCGATTACG TGTGGGTTTC GATTGCGGTG TTTTTTCGCG AATTTCCGCG
 421 TCCGTGTCAA AGTCCGGTGT TTATGTCACA GATCAGCTGA TCCACAGGGT ATTTAAACCA
 481 GTCGAGCCCG TCAAGAGGCC ACTCTTGAGT GCCAGCGAGT AGAGATTTCT CTGAGCTCCG
 541 CTCCCAGAGT GTGAGAAAAA TGAGACACCT GCGCCTCCTG CCTGGAACTG TGCCCTTGGA
 601 CATGGCCGCA TTATTGCTGG ATGACTTTGT GAGTACAGTA TTGGAGGATG AACTGCAACC
 661 AACTCCGTTC GAGCTGGGAC CCACACTTCA GGACCTCTAT GATTTGGAGG TAGATGCCCA
 721 GGAGGACGAC CCGAACGAAG ATGCTGTGAA TTTAATATTT CCAGAATCTC TGATTCTTCA
 781 GGCTGACATA GCCAGCGAAG CTCTACCTAC TCCACTTCAT ACTCCAACTC TGTCACCCAT
 841 ACCTGAATTG GAAGAGGAGG ACGAGTTAGA CCTCCGGTGT TATGAGGAAG GTTTTCCTCC
 901 CAGCGATTCA GAGGACGAAC AGGGTGAGCA GAGCATGGCT CTAATCTCAG ACTATGCTTG
 961 TGTGGTTGTG GAAGAGCATT TTGTGTTGGA CAATCCTGAG GTGCCCGGGC AAGGCTGTAA
1021 ATCCTGCCAG TACCACCGGG ATAAGACCGG AGACACGAAC GCCTCCTGTG CTCTGTGTTA
1081 CATGAAAAAG AACTTCAGCT TTATTTACAG TAAGTGGAGT GAATGTGAGA GAGGCTGAGT
1141 GCTTAAGACA TAACTGGGTG ATGCTTCAAC AGCTGTGCTA AGTGTGGTTT ATTTTGTTTC
1201 TAGGTCCGGT GTCAGAGGAT GGTCATCACC CTCAGAAGAA GACCACCCGT GTCCCCCTGA
1261 TCTGTCAGGC GAAACGCCCC TGCAAGTGCA CAGACCCACC CCAGTCAGAC CCAGTGGCGA
1321 GAGGCGAGCA GCTGTTGAAA AAATTGAGGA CTTGTTACAT GACATGGGTG GGGATGAACC
1381 TTTGGACCTG AGCTTGAAAC GTCCCAGGAA ACTAGGCGCA GCTGCGCTTA GTCATGTGTA
1441 AATAAAGTTG TACAATAAAA ATTATATGTG ACGCATGCAA GGTGTGGTTT ATGACTCATG
1501 GGCGGGGCTT AGTTCTATAT AAGTGGCAAC ACCTGGGCAC TGGAGCACAG ACCTTCAGGG
1561 AGTTCCTGAT GGATGTGTGG ACTATCCTTG CAGACTTTAG CAAGACACGC CGGCTTGTAG
1621 AGGATAGTTC AGACGGGTGC TCCGGGTTCT GGAGACACTG GTTTGGAACT CCTCTATCTC
1681 GCCTGGTGTA CACAGTTAAA AAGGATTATA ACGAGGAATT TGAAAATCTT TTTGCTGATT
1741 GCTCTGGCCT GCTAGATTCT CTGAATCTCG GCCACCAGTC CCTTTTCCAG GAAAGGGTAC
1801 TCCACAGCCT TGATTTTTCC AGCCCAGGGC GCACTACAGC CGGGGTTGCT TTTGTGGTTT
1861 TTCTGGTTGA CAAATGGAGC CAGAACACCC AACTGAGCAG GGGCTACATT CTGGACTTCG
1921 CAGCCATGCA CCTGTGGAGG GCATGGGTCA GGCAGCGGGG ACAGAGAATC TTGAACTACT
1981 GGCTTCTACA GCCAGCAGCT CCGGGTCTTC TTCGTCTACA CAGACAAACA TCCATGTTGG
2041 AGGAAGAAAT GAGGCAGGCC ATGGACGAGA ACCCGAGGAG CGGTCTGGAC CCTCCGTCGG
2101 AAGAGGAGTT GGATTGAATC AGGTATCCAG CCTGTACCCA GAGCTTAGCA AGGTGCTGAC
2161 ATCCATGGCC AGGGGAGTGA AGAGGGAGAG GAGCGATGGG GGCAATACCG GGATGATGAC
2221 CGAGCTGACG GCCAGTCTGA TGAATCGCAA GCGCCCAGAG CGCCTTACCT GGTACGAGCT
2281 ACAGCAGGAG TGCAGGGATG AGTTGGGCCT GATGCAGGAT AAATATGGCC TGGAGCAGAT
2341 AAAAACCCAT TGGTTGAACC CAGATGAGGA TTGGGAGGAG GCTATTAAGA AGTATGCCAA
2401 GATAGCCCTG CGCCCAGATT GCAAGTACAT AGTGACCAAG ACCGTGAATA TCAGACATGC
2461 TGCTACATCT CGGGGAACGG GGCAGAGGTG GTCATTGATA CCCTGGACAA GGCCGCCTTT
2521 AGGTGTTGCA TGATGGGAAT GAGAGCCGGA GTGATGAATA TGAATTCCAT GATCTTTATG
2581 AACATGAAGT TCAATGGAGA GAAGTTTAAT GGGGTGCTGT TCATGGCCAA CAGCCACATG
2641 ACCCTGCATG GCTGCGACTT TTTCGGCTTT AACAATATGT GCGCAGAGGT CTGGGGCGCT
2701 TCCAAGATCA GGGGATGTAA GTTTTATGGC TGCTGGATGG GCGTGGTCGG AAGACCCAAG
2761 AGCGAGATGT CTGTGAAGCA GTGTGTGTTT GAGAAATGCT ACCTGGGAGT CTCTACCGAG
2821 GGCAATGCTA GAGTGAGGCA CTGCTCTTCC CTGGAGACGG GCTGCTTCTG CCTGGTGAAG
2881 GGCACAGCCT CTCTGAAGCA TAATATGGTG AAGGGCTGCA CGGATGAGCG CATGTACAAC
2941 ATGCTGACTG CGACTCGGGG GTCTGTCATA TCCTGAAGAA CATCCATGTG ACCTCCCACC
3001 CCAGAAAGAA GTGGCCAGTG TTTGAGAATA ACATGCTGAT CAAGTGCCAC ATGCACCTGG
3061 GCGCCAGAAG GGGCACCTTC CAGCCGTACC AGTGCAACTT TAGCCAGACC AAGCTGCTGT
3121 TGGAGAACGA TGCCTTCTCC AGGGTGAACC TGAACGGCAT CTTTGACATG GATGTCTCGG
3181 TGTACAAGAT CCTGAGATAC GATGAGACCA AGTCCAGGGT GCGCGCTTGC GAGTGCGGGG
3241 GCAGACACAC CAGGATGCAG CCAGTGGCCC TGGATGTGAC CGAGGAGCTG AGACCAGACC
3301 ACCTGGTGAT GGCCTGTACC GGGACCGAGT TCAGCTCCAG TGGGGAGGAC ACAGATTAGA
3361 GGTAGGTTTG AGTAGTGGGC GTGGCTAAGG TGACTATAAA GGCGGGTGTC TTACGAGGGT
```

3421 CTTTTTGCTT TTCTGCAGAC ATCATGAACG GGACCGGCGG GGCCTTCGAA GGGGGGCTTT
3481 TTAGCCCTTA TTTGACAACC CGCCTGCCAG GATGGGCCGG AGTTCGTCAG AATGTGATGG
3541 GATCGACGGT GGACGGGCGC CCAGTGCTTC CAGCAAATTC CTCGACCATG ACCTACGCGA
3601 CCGTGGGGAA CTCGTCGCTT GACAGCACCG CCGCAGCCGC GGCAGCCGCA GCCGCCATGA
3661 CAGCGACGAG ACTGGCCTCG AGCTACATGC CCAGCAGCAG CAGTAGCCCC TCTGTGCCCA
3721 GTTCCATCAT CGCCGAGGAG AACTGCTGGC CCTGCTGGCC GAGCTGGAAG CCCTGAGCCG
3781 CCAGCTGGCC GCCCTGACCC AGCAGGTGTC CGAGCTCCGC GAACAGCAGC AGCAAAATAA
3841 ATGATTCAAT AAACACATAT TCTGATTCAA ACAGCAAAGC ATCTTTATTA TTTATTTTTT
3901 CGCGCGCGGT AGGCCCTGGT CCACCTCTCC CGATCATTGA GAGTGCGGTG GATTTTTTCC
3961 AAGACCCGGT AGAGGTGGGA TTGGATGTTG AGGTACATGG GCATGAGCCC GTCCCGGGGG
4021 TGGAGGTAGC ACCACTGCAT GGCCTCGTGC TCTGGGGTCG TGTTGTAGAT GATCCAGTCA
4081 TAGCAGGGGC GCTGGGCGTG GTGCTGGATG ATGTCCTTGA GGAGGAGACT GATGGCCACG
4141 GGGAGCCCCT TGGTGTAGGT GTTGGCAAAG CGGTTGAGCT GGGAGGGATG CATGCGGGGG
4201 GAGATGATGT GCAGTTTGGC CTGGATCTTG AGGTTGGCGA TGTTGCCACC CAGATCCCGC
4261 CGGGGGTTCA TGTTGTGCAG GACCACCAGG ACGGTGTAGC CCGTGCACTT GGGGAACTTA
4321 TCATGCAACT TGGAAGGGAA TGCGTGGAAG AATTTGGAGA CGCCCTTGTG CCCGCCCAGG
4381 TTTTCCATGC ACTCATCCAT GATGATGGCG ATGGGCCCGT GGGCTGCGGC TTTGGCAAAG
4441 ACGTTTCTGG GGTCAGAGAC ATCATAATTA TGCTCCTGGG TGAGATCATC ATAAGACATT
4501 TTAATGAATT TTGGGCGGAG GGTGCCAGAT TGGGGGACGA TGGTTTCCCT CGGGCCCCGG
4561 GGCGAAGTTC CCCTCGCAGA TCTGCATCTC CCAGGCTTTC ATCTCGGAGG GGGGGATCAT
4621 GTCCACCTGC GGGGCGATGA AAAAAACGGT TTCCGGGGCG GGGGTGATGA GCTGCGAGGA
4681 GAGCAGGTTT CTCAACAGCT GGGACTTGCC GCACCCGGTC GGGCCGTAGA TGACCCCGAT
4741 GACGGGTTGC AGGTGGTAGT TCAAGGACAT GCAGCTGCCG TCGTCCCGGA GGAGGGGGGC
4801 CACCTCGTTG AGCATGTCTC TAACTTGGAG GTTTTCCCGG ACGAGCTCGC CGAGGAGGCG
4861 GTCCCCGCCC AGCGAGAGGA GCTCTTGCAG GGAAGCAAAG TTTTTCAGGG GCTTGAGTCC
4921 GTCGGCCATG GGCATCTTGG CGAGGGTCTG CGAGAGGAGT TCGAGACGTC CCAGAGCTCG
4981 GTGACGTGCT CTACGGCATC TCGATCCAGC AGACTTCCTC GTTTCGGGGG TTGGGACGAC
5041 TGCGACTGTA GGGCACGAGA CGATGGGCGT CCAGCGCGGC CAGCGTCATG TCCTTCCAGG
5101 GTCTCAGGGT CCGCGTGAGG GTGGTCTCCG TCACGGTGAA GGGGTGGGCC CCTGGCTGGG
5161 CGCTTGCAAG GGTGCGCTTG AGACTCATCC TGCTGGTGCT GAAACGGGCA CGGTCTTCGC
5221 CCTGCGCGTC GGCGAGATAG CAGTTGACCA TGAGCTCGTA GTTGAGGGCC TCGGCGGCGT
5281 GGCCCTTGGC GCGGAGCTTG CCCTTGGAAG AGCGTCCGCA GGCGGGACAG AGGAGGGATT
5341 GCAGGGCGTA GAGCTTGGGC GCAAGAAAGA CCGACTCGGG AGCAAAAGCG TCCGCTCCGC
5401 AGTGGGCGCA GACGGTCTCG CACTCGACGA GCCAGGTGAG CTCGGGCTGC TCGGGGTCAA
5461 AAACCAGTTT TCCCCCGTTC TTTTTGATGC GCTTCTTACC TCGCGTCTCC ATGAGTCTGT
5521 GTCCGCGCTC GGTGACAAAC AGGCTGTCGG TGTCCCCGTA GACGGACTTG ATTGGCCTGT
5581 CCTGCAGGGG CGTCCCGCGG TCCTCCTCGT AGAGAAACTC GGACCACTCT GAGACAAAGG
5641 CGCGCGTCCA CGCCAAGACA AAGGAGGCCA CGTGCGAGGG GTAGCGGTCG TTGTCCACCA
5701 GGGGGTCCAC CTTTTCCACC GTGTGCAGAC ACATGTCCCC TTCCTCCGCA TCCAAGAAGG
5761 TGATTGGCTT GTAGGTGTAG GCCACGTGAC CAGGGGTCCC CGACGGGGGG GTATAAAAGG
5821 GGGCGGGTCT GTGCTCGTCC TCACTCTCTT CCGCGTCGCT GTCCACGAGC GCCAGCTGTT
5881 GGGGTAGGTA TTCCCTCTCG AGAGCGGGCA TGACCTCGGC ACTCAGGTTG TCAGTTTCTA
5941 GAAACGAGGA GGATTTGATG TTGGCTTGCC CTGCCGCAAT GCTTTTTAGG AGACTTTCAT
6001 CCATCTGGTC AGAAAAGACT ATTTTTTTAT TGTCAAGCTT GGTGGCAAAG GAGCCATAGA
6061 GGGCGTTGGA GAGAAGCTTG GCGATGGATC TCATGGTCTG ATTTTTGTCA CGGTCGGCGC
6121 GCTCCTTGGC CGCGATGTTG AGCTGGACAT ATTCGCGCGC GACACACTTC CATTCGGGAA
6181 AGACGGTGGT GCGCTCGTCG GGCACGATCC TGACGCGCCA GCCGCGGTTA TGCAGGGTGA
6241 CCAGGTCCAC GCTGGTGGCC ACCTCGCCGC GCAGGGGCTC GTTAGTCCAG CAGAGTCTGC
6301 CGCCCTTGCG CGAGCAGAAC GGGGGCAGCA CATCAAGCAG ATGCTCGTCA GGGGGGTCCG
6361 CATCGATGGT GAAGATGCCG GGACAGAGTT TCTTGTCAAA ATAGTCTATT TTTGAGGATG
6421 CATCATCCAA GGCCATCTGC CACTCGCGGG CGGCCATTGC TCGCTCGTAG GGGTTGAGGG
6481 GCGGACCCCA CGGCATGGGA TGCGTGAGGG CGGAGGCGTA CATGCCGCAA ATGTCGTAAA
6541 CATAGATGGG CTCCGAGAAG ATGCCGATGT TGGTGGGATA ACAGCGCCCC CCGCGGATGC
6601 TGGCGCGCAC GTATTCATAC AACTCGTGCG AGGGGCCAAG AAGGCCGGGG CCGAAATTGG
6661 TGCGCTGGGG CTGCTCGGCG CGGAAAACAA TCTGGCGAAA GATGGCGTGC GAGTTGGAGG
6721 AGATGGTGGG CCGTTGGAAG ATGTTAAAGT GGGCGTGGGG CAAGCGGACC GAGTCGCGGA
6781 TGAAGTGCGC GTAGGAGTCT TGCAGCTTGG CGACGAACTC GGCGGTGACG AGAACGTCCA

6841 TGGCGCAGTA GTCCAGCGTT TCGCGGATGA TGTCATAACC CGCCTCTCCT TTCTTCTCCC
6901 ACAGCTCGCG GTTGAGGGCG TATTCCTCGT CATCCTTCCA GTACTCCCGG AGCGGGAATC
6961 CTCGATCGTC CGCACGGTAA GAGCCCAGCA TGTAGAAATG GTTCACGGCC TTGTAGGGAC
7021 AGCAGCCCTT CTCCACGGGG AGGGCGTAAG CTTGTGCGGC CTTGCGGAGC GAGGTGTGCG
7081 TCAGGGCGAA GGTGTCCCTG ACCATGACTT TCAAGAACTG GTACTTGAAA TCCGAGTCGT
7141 CGCAGCCGCC GTGCTCCCAT AGCTCGAAAT CGGTGCGCTT CTTCGAGAGG GGGTTAGGCA
7201 GAGCGAAAGT GACGTCATTG AAGAGAATCT TGCCTGCTCG CGGCATGAAA TTGCGGGTGA
7261 TGCGGAAAGG GCCCGGGACG GAGGCTCGGT TGTTGATGAC CTGGGCGGCG AGGACGATCT
7321 CGTCGAAGCC GTTGATGTTG TGCCCGACGA TGTAGAGTTC CATGAATCGC GGGCGGCCTT
7381 TGATGTGCGG CAGCTTTTTG AGCTCCTCGT AGGTGAGGTC CTCGGGGCAT TGCAGGCCGT
7441 GCTGCTCGAG CGCCCATTCC TGGAGATGTG GGTTGGCTTG CATGAAGGAA GCCCAGAGCT
7501 CGCGGGCCAT GAGGGTCTGG AGCTCGTCGC GAAAGAGGCG GAACTGCTGG CCCACGGCCA
7561 TCTTTTCGGG TGTGACGCAG TAGAAGGTGA GGGGGTCCCG CTCCCAGCGA TCCCAGCGTA
7621 AGCGCGCGGC TAGATCGCGA GCAAGGGCGA CCAGCTCTGG GTCCCCCGAG AATTTCATGA
7681 CCAGCATGAA GGGGACGAGC TGCTTGCCGA AGGACCCCAT CCAGGTGTAG GTTTCTACAT
7741 CGTAGGTGAC AAAGAGCCGC TCCGTGCGAG GATGAGAGCC GATTGGGAAG AACTGGATTT
7801 CCTGCCACCA GTTGGACGAG TGGCTGTTGA TGTGATGAAA GTAGAAATCC CGCCGGCGAA
7861 CCGAGCACTC GTGCTGATGC TTGTAAAAGC GTCCGCAGTA CTCGCAGCGC TGCACGGGCT
7921 GTACCTCATC CACGAGATAC ACAGCGCGTC CCTTGAGGAG GAACTTCAGG AGTGGCGGCC
7981 CTGGCTGGTG GTTTTCATGT TCGCCTGCGT GGGACTCACC CTGGGGCTCC TCGAGGACGG
8041 AGAGGCTGAC GAGCCCGCGC GGGAGCCAGG TCCAGATCTC GGCGCGGCGG GGGCGGAGAG
8101 CGAAGACGAG GGCGCGCAGT TGGGAGCTGT CCATGGTGTC GCGGAGATCC AGGTCCGGGG
8161 GCAGGGTTCT GAGGTTGACC TCGTAGAGGC GGGTGAGGGC GTGCTTGAGA TGCAGATGGT
8221 ACTTGATTTC TACGGGTGAG TTGGTGGCCG TGTCCACGCA TTGCATGAGC CCGTAGCTGC
8281 GCGGGGCCAC GACCGTGCCG CGGTGCGCTT TTAGAAGCGG TGTCGCGGAC GCGCTCCCGG
8341 CGGCAGCGGC GGTTCCGGCC CCGCGGGCAG GGGCGGCAGA GGCACGTCGG CGTGGCGCTC
8401 GGGCAGGTCC CGGTGTTGCG CCCTGAGAGC GCTGGCGTGC GCGACGACGC GGCGGTTGAC
8461 ATCCTGGATC TGCCGCCTCT GCGTGAAGAC CACTGGCCCC GTGACTTTGA ACCTGAAAGA
8521 CAGTTCAACA GAATCAATCT CGGCGTCATT GACGGCGGGC TGACGCAGGA TCTCTTGCAC
8581 GTCGCCCGAG TTGTCCTGGT AGGCGATCTC GGACATGAAC TGCTCGATCT CCTCCTCCTG
8641 GAGATCGCCG CGACCCGCGC GCTCCACGGT GGCGGCGAGG TCATTCGAGA TGCGACCCAT
8701 GAGCTGCGAG AAGGCGCCCA GGCCGCTCTC GTTCCAGACG CGGCTGTAGA CCACGTCCCC
8761 GTCGGCGTCG CGCGCGCGCA TGACCACCTG CGCGAGGTTG AGCTCCACGT GCCGCGCGAA
8821 GACGGCGTAG TTGCGCAGGC GCTGGAAGAG GTAGTTGAGG GTGGTGGCGA TGTGCTCGGT
8881 GACGAAGAAG TACATGATCC AGCGGCGCAG GGGCATCTCG CTGATGTCGC CGATGGCCTC
8941 CAGCCTTTCC ATGGCCTCGT AGAAATCCAC GGCGAAGTTG AAAAACTGGG CGTTGCGGGC
9001 CGAGACCGTG AGCTCGTCTT CCAGGAGCCT GATGAGCTCG GCGATGGTGG CGCGCACCTC
9061 GCGCTCGAAA TCCCCGGGGG CCTCGTCCTC TTCCTCTTCT TCCATGACAA CCTCTTCTAT
9121 TTCTTCCTCT GGGGGCGGTG GTGGTGGCGG GGCCCGACGA CGACGGCGAC GCACCGGGAG
9181 ACGGTCGACG AAGCGCTCGA TCATCTCCCC GCGGCGGCGA CGCATGGTTT CGGTGACGGC
9241 GCGACCCCGT TCGCGAGGAC GCAGCGTGAA GACGCCGCCG GTCATCTCCC GGTAATGGGG
9301 CGGGTCCCCG TTGGGCAGCG AGAGGGCGCT GACGATGCAT CTTATCAATT GCGGTGTAGG
9361 GGACGTGAGC GCGTCGAGAT CGACCGGATC GGAGAATCTT TCGAGGAAAG CGTCTAGCCA
9421 ATCGCAGTCG CAAGGTAAGC TCAAACACGT AGCAGCCCTG TGGACGCTGT TAGAATTGCG
9481 GTTGCTAATG ATGTAATTGA AGTAGGCGTT TTTGAGGCGG CGGATGGTGG CGAGGAGGAC
9541 CAGGTCCTTG GGTCCCGCTT GCTGGATGCG GAGCCGCTCG GCCATGCCCC AGGCCTGGCC
9601 CTGACACCGG CTTAGGTTCT TGTAGTAGTC ATGCATGAGC CTCTCGATGT CATCACTGGC
9661 GGAGGCGGAG TCTTCCATGC GGGTGACCCC GACGCCCCTG AGCGGCTGCA CGAGCGCCAG
9721 GTCGGCGACG ACGCGCTCGG CGAGGATGGC CTGTTGCACG CGGGTGAGGG TGTCCTGGAA
9781 GTCGTCCATG TCGACGAAGC GGTGGTAGGC CCCTGTGTTG ATGGTGTAAG TGCAGTTGGC
9841 CATGAGCGAC CAGTTGACGG TCTGCAGGCC GGGCTGCACG ACCTCGGAGT ACCTGAGCCG
9901 CGAGAAGGCG CGCGAGTCGA AGACGTAGTC GTTGCAGGTG CGCACAAGGT ACTGGTATCC
9961 GACTAGGAAG TGCGGCGGCG GCTGGCGGTA GAGCGGCCAG CGCTGGGTGG CCGGCGCGCC
10021 CGGGGCCAGG TCCTCGAGCA TGAGGCGGTG GTAGCCGTAG AGGTAGCGGG ACATCCAGGT
10081 GATGCCGGCA GCGGTGGTGG AGGCGCGCGG GAACTCGCGG ACGCGGTTCC AGATGTTGCG
10141 CAGCGGCAGG AAATAGTCCA TGGTCGGCAC GGTCTGGCCG GTGAGACGCG CGCAGTCATT
10201 GACGCTCTAG AGGCAAAAAC GAAAGCGGTT GAGCGGGCTC TTCCTCCGTA GCCTGGCGGA

10261 ACGCAAACGG GTTAGGCCGC GCGTGTACCC CGGTTCGAGT CCCCTCGAAT CAGGCTGGAG
10321 CCGCGACTAA CGTGGTATTG GCACTCCCGT CTCGACCCGA GCCCGATAGC CGCCAGGATA
10381 CGCGGGAAGA GCCCTTTTTG CCGGCCGARG GGAGTCGCTA GACTTGAAAG CGGCCGAAAA
10441 CCCCGCCGGG TAGTGGCTCG CGCCCGTAGT CTGGAGAAGC ATCGCCAGGG TTGAGTCGCG
10501 GCAGAACCCG GTTCGCGGAC GGCCGCGGCG AGCGGGACTT GGTCACCCCG CCGATTTAAA
10561 GACCCACAGC CAGCCGACTT CTCCAGTTAC GGGAGCGAGC CCCCTTTTTT CTTTTTGCCA
10621 GATGCATCCC GTCCTGCGCC AAATGCGTCC CACCCCCCCG GCGACCACCG CGACCGCGGC
10681 CGTAGCAGGC GCCGGCGCTA GCCAGCCACA GCCACAGACA GAGATGGACT TGGAAGAGGG
10741 CGAAGGGCTG GCGAGACTGG GGGCGCCTTC CCCGGAGCGA CACCCCCGCG TGCAGCTGCA
10801 GAAGGACGTG CGCCCGGCGT ACGTGCCTGC GCAAAACCTG TTCAGGGACC GCAGCGGGGA
10861 GGAGCCCGAG GAGATGCGCG ACTGCCGGTT TCGGGCGGGC AGGGAGCTGC GCGAGGGCCT
10921 GGACCGCCAG CGCGTGCTGC GCGACGAGGA TTTCGAGCCG AACGAGCAGA CGGGGATCAG
10981 CCCCGCGCGC GCGCACGTGG CGGCGGCCAA CCTGGTGACG GCCTACGAGC AGACGGTGAA
11041 GCAGGAGCGC AACTTCCAAA AGAGTTTCAA CAACCATGTG CGCACCCTGA TCGCGCGCGA
11101 GGAGGTGGCC CTGGGCCTGA TGCACCTGTG GGACCTGGCG GAGGCCATCG TGCAGAACCC
11161 GGACAGCAAG CCTCTGACGG CGCAGCTGTT CCTGGTGGTA CAGCACAGCA GGGACAACGA
11221 GGCGTTCAGG GAGGCGCTGC TAAACATCGC CGAGCCCGAG GGTCGCTGGC TGCTGGAGCT
11281 GATCAACATC TTGCAGAGCA TCGTAGTTCA GGAGCGCAGC CTGAGCTTGG CCGAGAAGGT
11341 GGCGGCAATC AACTACTCGG TGCTTAGCCT GGGCAAGTTT TACGCGCGCA AGATTTACAA
11401 GACGCCGTAC GTGCCCATAG ACAAGGAGGT GAAGATAGAC AGCTTTTACA TGCGCATGGC
11461 GCTCAAGGTG CTGACGCTGA GCGACGACCT GGGCGTGTAC CGCAACGACC GCATCCACAA
11521 GGCCGTGAGC GCGAGCCGGC GGCGCGAGCT GAGCGACCGC GAGCTGATGC TGAGCCTGCG
11581 CCGGGCGCTG GTAGGGGGCG CCGCCGGCGG CGAGGAGTCY TACTTCGACA TGGGGGCGGA
11641 CCTGCATTGG CAGCCGAGCC GGCGCGCCTT GGAGGCCGCC TACGGTCCAG AGGACTTGGA
11701 TGAGGAAGAG GAAGAGGAGG AGGATGCACC CGCTGCGGGG TACTGACGCC TCCGTGATGT
11761 GTTTTTAGAT GCAGCAAGCC CCGGACCCCG CCATAAGGGC GGCGCTGCAA AGCCAGCCGT
11821 CCGGTCTAGC ATCGGACGAC TGGGAGGCTG CGATGCAACG CATCATGGCC CTGACGACCC
11881 GCAACCCCGA GTCCTTTAGA CAACAGCCGC AGGCCAACAG ACTCTCGGCC ATTCTGGAGG
11941 CGGTGGTCCC TTCTCGGACC AACCCCACGC ACGAGAAGGT GCTGGCGATC GTGAACGCGC
12001 TGGCGGAGAA CAAGGCCATC CGTCCCGACG AGGCCGGGCT AGTGTACAAC GCCCTGCTGG
12061 AGCGCGTAGG CCGCTACAAC AGCACAAACG TGCAGTCCAA CCTGGACCGG CTGGTGACGG
12121 ACGTGCGCGA AGCCGTGGCG CAGCGCGAGC GGTTCAAGAA CGAGGGCCTG GGCTCGCTGG
12181 TGGCGCTGAA CGCCTTCCTG GCGACGCAGC CGGCGAACGT GCCGCGCGGG CAGGATGATT
12241 ACACCAACTT TATCAGCGCG CTGCGGCTGA TGGTGACCGA GGTGCCCCAG AGCGAGGTGT
12301 ACCAGTCGGG CCCGGACTAC TTTTTCCAAA CTAGCAGACA GGGCCTGCAA ACGGTGAACC
12361 TGAGCCAGGC TTTCAAGAAC CTGCGCGGGC TGTGGGGCGT GCAGGCGCCC GTGGGCGACC
12421 GGTCGACGGT GAGCAGCTTG CTGACGCCCA ACTCGCGGCT GCTGCTGCTG CTGATCGCGC
12481 CCTTCACCGA CAGTGGCAGC GTAAACCGCA ACTCGTACCT GGGTCACCTG CTAACGCTGT
12541 ACCGCGAGGC CATAGGCCAG GCGCAGGTGG ACGAGCAGAC CTTCCAGGAG ATCACTAGCG
12601 TGAGCCGCGC GCTGGGGCAG AACGACACCG ACAGTCTGAG GGCCACCCTG AACTTCTTGC
12661 TGACCAATAG ACAGCAGAAG ATCCCGGCGC AGTACGCGCT GTCGGCCGAG GAGGAGCGCA
12721 TCCTGAGATA TGTGCAGCAG AGCGTAGGGC TTTTCCTGAT GCAGGAGGGG GCCACTCCCA
12781 GCGCCGCGCT GGACATGACC GCGCGCAACA TGGAACCTAG CATGTACGCC GCCAACCGGC
12841 CGTTTATCAA TAAGCTAATG GACTACCTGC ATCGCGCGGC GTCCATGAAC TCGGACTACT
12901 TTACCAATGC CATTTTGAAC CCGCACTGGC TTCCGCCGCC GGGGTTCTAT ACGGGCGAGT
12961 ACGACATGCC CGACCCCAAC GACGGGTTTT TGTGGGACGA CGTGGACAGC GCGGTGTTTT
13021 CACCGACCTT GCAAAAGCGC CAGGAGGCGG TGCGCACGCC CGCGAGCGAG GGCGCGGTGG
13081 GTCGGAGCCC CTTTCCTAGC TTAGGGAGTT TGCATAGCTT GCCGGGCTCT GTGAACAGCG
13141 GCAGGGTGAG CCGGCCGCGC TTGCTGGGCG AGGACGAGTA CCTGAACGAC TCGCTGCTGC
13201 AGCCGCCGCG GGTCAAGAAC GCCATGGCCA ATAACGGGAT AGAGAGTCTG GTGGACAAAC
13261 TGAACCGCTG GAAGACCTAC GCTCAGGACC ATAGGGAGCC TGCGCCCGCG CCGCGGCGAC
13321 AGCGCCACGA CCGGCAGCGG GGCCTGGTGT GGGACGACGA GGACTCGGCC GACGATAGCA
13381 GCGTGTTGGA CTTGGGCGGG AGCGGTGGGG TCAACCCGAT ATCGCGCATC CTGCAGCCCA
13441 AACTGGGGCG ACGGATGTTT TGAATGCAAA ATAAAACTCA CCAAGGCCAT AGCGTGCGTT
13501 CTCTTCCTTG TTAGAGATGA GGCGTGCGGT GGTGTCTTCC TCTCCTCCTC CCTCGTACGA
13561 GAGCGTGATG GCGCAGGCGA CCCTGGAGGT TCCGTTTGTG CCTCCGCGGT ATATGGCTCC
13621 TACGGAGGGC AGAAACAGCA TTCGTTACTC GGAGCTGGCT CCGTTGTACG ACACCACTCG

13681 CGTGTACTTG GTGGACAACA AGTCGGCGGA CATCGCTTCC CTGAACTATC AAAACGACCA
13741 CAGCAACTTC CTGACCACGG TGGTGCAGAA CAACGATTTC ACCCCCGCCG AGGCTAGCAC
13801 GCAGACGATA AATTTTGACG AGCGGTCGCG GTGGGGCGGT GATCTGAAGA CCATTCTGCA
13861 CACCAACATG CCCAATGTGA ACGAGTACAT GTTCACCAGC AAGTTTAAGG CGCGGGTGAT
13921 GGTGGCTAGA AAACACCCAC AGGGGGTAGA AGCAACAGAT TTAAGCAAGG ATATCTTAGA
13981 GTATGAGTGG TTTGAGTTTA CCCTGCCCGA GGGCAACTTT TCCGAGACCA TGACCATAGA
14041 CCTGATGAAC AACGCCATCT TGGAAAACTA CTTGCAAGTG GGGCGGCAAA ATGGCGTGCT
14101 GGAGAGCGAT ATTGGAGTCA AGTTTGACAG CAGAAATTTC AAGCTGGGCT GGGACCCTGT
14161 GACCAAGCTG GTGATGCCAG GGGTCTACAC CTACGAGGCC TTTCACCCGG ACGTGGTGCT
14221 GCTGCCGGGC TGCGGGGTGG ACTTCACAGA GAGCCGCCTG AGCAACCTCC TGGGCATTCG
14281 CAAGAAGCAA CCTTTCCAAG AGGGCTTCAG AATCATGTAT GAGGATCTAG AAGGGGGCAA
14341 CATCCCCGCC CTGCTGGATG TGCCCAAGTA CTTGGAAAGC AAGAAGAAGT TAGAGGAGGC
14401 ATTGGAGAAT GCTGCTAAAG CTAATGGTCC TGCAAGAGGA GACAGTAGCG TCTCAAGAGA
14461 GGTTGAAAAG GCAGCTGAAA AAGAACTTGT TATTGAGCCC ATCAAGCAAG ATGATACCAA
14521 GAGAAGTTAC AACCTCATCG AGGGAACCAT GGACACGCTG TACCGCAGCT GGTACCTGTC
14581 CTATACCTAC CGGGACCCTG AGAACGGGGT GCAGTCGTGG ACGCTGCTCA CCACCCCGGA
14641 CGTCACCTGC GGCGCGGAGC AAGTCTACTG GTCGCTGCCG GACCTCATGC AAGACCCCGT
14701 CACCTTCCGT TCTACCCAGC AAGTCAGCAA CTACCCCGTG GTCGGCGCCG AGCTCATGCC
14761 CTTCCGCGCC AAGAGCTTTT ACAACGACCT CGCCGTCTAC TCCCAGCTCA TCCGCAGCTA
14821 CACCTCCCTC ACCCACGTCT TCAACCGCTT CCCCGACAAC CAGATCCTCT GCCGTCCGCC
14881 CGCGCCCACC ATCACCACCG TCAGTGAAAA CGTGCCTGCT CTCACAGATC ACGGGACGCT
14941 ACCGCTGCGC AGCAGTATCC GCGGAGTCCA GCGAGTGACC GTCACTGACG CCCGTCGCCG
15001 CACCTGTCCC TACGTCTACA AGGCCCTGGG CATAGTCGCG CCGCGTGTGC TTTCCAGTCG
15061 CACCTTCTAA AAAATGTCTA TTCTCATCTC GCCCAGCAAT AACACCGGCT GGGGTATTAC
15121 TAGGCCCAGC AGCATGTACG GAGGAGCCAA GAAACGTCCC AGCAGCACCC CGTCCGCGTC
15181 CGCGGCCACT TCCGCGCTCC GTGGGGCGCT TACAAGCGCG GCGGACTGC CACCGCCGCC
15241 GCCGTGCGCA CCACCGTCGA CGACGTCATC GACTCGGTGG TCGCCGACGC GCGCAACTAT
15301 ACTCCCGCCC CTTCGACCGT GGACGCGGTT CATTGACAGC GTGGTGGCGA CGCGGCGGCG
15361 ATATGCCAGA CGCAAGAGCC GGCGGGCGGA CGGATCGCCC AGGCGCCATT CGGAGCACGC
15421 CCGCCATGGG GCGCCGCCCG AGCTCTGCTG CGCCGCGCCA GACGCACGGG CCGCCGGGCC
15481 ATGATGCGAG CCGCGCGCCG CGCCGCCACT GCACCCCCCG CAGGCAGGAC TCGCAGACGA
15541 GCGGCCGCCG CCGCCGCCGC GGCCATCTCT AGCATGACCA GACCCAGGCG CGGAAACGTG
15601 TACTGGGTGC GCGACTCCGT CACGGGCGTG CGCGTGCCCG TGCGCACCCG TCCTCCTCGT
15661 CCCTGATCTA ATGCTTGTGT CCTCCCCCGC AAGCGACGAT GTCAAAGCGC ATCTACAAGA
15721 GAGATGCTCC AGGTCGTCGC CCCGGAGATT TACGGACCAC CCCAGGCGGA CCAGAAACCC
15781 CGCAAAATCA AGCGGGTTAA AAAAAAGGAT GAGGTGGACG AGGGGGCAGT AGAGTTTGTG
15841 CGCGAGTTCG CTCCGCGGCG GCGCGTAAAT TGGAAGGGGC GCAGGTGCAC GCGTGTTGCG
15901 GCCCGGCACG GCGGTGGTGT TCACGCCCGG CGAGCGGTCC TCGGTCAGGA GCAAGCGTAG
15961 CTATGACGAG GTGTACGGCG ACGACGACAT CCTGGACCAG GCGGCAGAGC GGGCGGGCGA
16021 GTTTGCCTAC GGGAAGCGGT CGCGCGAAGA GGAGCTGATC TCGCTGCCGC TGGACGAGAG
16081 CAATCCCACG CCGAGCCTGA AGCCCGTGAC CTGCAGCAGG TGCTGCCCCA GGCGGTGCTG
16141 CTGCCGAGCC GCGGGATCAA GCGCGAGGGC GAGAACATGT ACCCGACCAT GCAGATCATG
16201 GTGCCCAAGC GCCGGCGCGT GGAGGAAGTG CTGGACACCG TGAAAATGGA TGTGGAGCCC
16261 GAGGTCAAGG TGCGCCCCAT CAAGCAGGTG GCGCCGGGCC TGGGCGTGCA GACCGTGGAC
16321 ATTCAGATCC CCACCGACAT GGATGTCGAC AAAAAACCCT CGACCAGCAT CGAGGTGCAG
16381 ACCGACCCCT GGCTCCCAGC CTCCACCGCT ACCGCTTCCA CTTCTACCGT CGCCACGGTC
16441 ACCGAGCCTC CCAGGAGGCG AAGATGGGGC CCCGCCAACC GGCTGATGCC CAACTACGTG
16501 TTGCATCCTT CCATTATCCC GACGCCGGGC TACCGCGGCA CCCGGTACTA CGCCAGCCGC
16561 AGGCGCCCAG CCAGCAAACG CCGCCGCCGC ACCGCCACCC GCCGCCGTCT GCCCCCCGCC
16621 CGCGTGCGCC GCGTAACCAA CGCGCCGGGG CCGCTCGCTC GTTCTGCCCA CCGTGCGCTA
16681 CCACCCCAGC ATCCTTTAAT CCGTGTGCTG TGATACTGTT GCAGAGAGAT GGCTCTCACT
16741 TGCCGCCTGC GCATCCCCGT TCCGAATTAC CGAGGAAGAT CCCGCCGCAG GAGAGGCATG
16801 GCAGGCAGCG GCCTGAACCG CCGCCGGCGG CGGGCCATGC GCAGGCGCCT GAGTGGCGGC
16861 TTTCTGCCCG CGCTCATCCC CATAATCGCG GCGGCCATCG GCACGATCCC GGGCATAGCT
16921 TCCGTTGCGC TGCAGGCGTC GCAGCGCCGT TGATGTGCGA ATAAAGCCTC TTTAGACTCT
16981 GACACACCTG GTCCTGTATA TTTTTAGAAT GGAAGACATC AATTTTGCGT CCCTGGCTCC
17041 GCGGCACGGC ACGCGGCCGT TCATGGGCAC CTGGAACGAG ATCGGCACCA GCCAGCTGAA

17101 CGGGGGCGCC TTCAATTGGA GCAGTGTCTG GAGCGGGCTT AAAAATTTCG GCTCGACGCT
17161 CCGGACCTAT GGGAACAAGG CCTGGAATAG TAGCACGGGG CAGTTGTTGA GGGAAAAGCT
17221 CAAAGACCAG AACTTCCAGC AGAAGGTGGT GGACGGCCTG GCCTCGGGCA TTAACGGGGT
17281 GGTGGACATC GCGAACCAGG CAGTGCAGCG CGAGATAAAC AGCCGTCTGG ACCCGCGGCC
17341 GCCCACGGTG GTGGAGATGG AAGATGCAAC TCTTCCGCCG CCGAAGGGCG AGAAGCGGCC
17401 GCGGCCAGAT GCGGAGGAGA CGATCCTGCA GGTGGACGAG CCGCCTTCGT ACGAGGAGGC
17461 CGTGAAGGCC GGCATGCCCA CCACGCGCAT CATCGCGCCA CTGGCCACGG GTGTAATGAA
17521 ACCCGCCACC CTTGACCTGC CTCCACCACC CACGCCCGCT CCACCGAAGG CAGCTCCGGT
17581 TGTGCAGCCC CCTCCGGTGG CGACCGCCGT GCGCCGCGTC CCGCCCGCCC GCCAGGCCCA
17641 GAACTGGCAG AGCACGCTGC ACAGTATTGT GGGCCTGGGA GTGAAAAGTC TGAAGCGCCG
17701 CCGATGCTAT TGAGAGAGAG GAAGGAGGAC ACTAAAGGGA GAGCTTAACT TGTATGTGCC
17761 TTACCGCCAG AGAACGCGCG AAGATGGCCA CCCCCTCGAT GATGCCGCAG TGGGCGTACA
17821 TGCACATCGC CGGGCAGGAC GCCTCGGAGT ACCTGAGCCC GGGTCTGGTG CAGTTTGCCC
17881 GCGCCACCGA CACGTACTTC AGCCTGGGCA ACAAGTTTAG GAACCCCACG GTGGCCCCGA
17941 CCCACGATGT GACCACGGAC CGGTCCCAGC GTCTGACGCT GCGCTTTGTG CCCGTGGATC
18001 GCGAGGACAC CAGTACTCGT ACAAGGCGCG CTTCACTCTG GCCGTGGGCG ACAACCGGGT
18061 GCTAGACATG GCCAGCACGT ACTTTGACAT CCGCGGCGTC CTGGACCGCG GTCCCAGTTT
18121 CAAACCCTAC TCGGGCACGG CTTACAACAG CCTTGCCCCC AAGGGCGCTC CCAATCCCAG
18181 TCAGTGGGTT GCCAAAGAAA ATGGTCAGGG AACTGATAAG ACACATACTT ATGGCTCAGC
18241 TGCCATGGGA GGAAGCAACA TCACCATTGA AGGTTTAGTA ATTGGAACTG ATGAAAAAGC
18301 TGAGGATGGC AAAAAAGATA TTTTTGCAAA TAAACTTTAT CAGCCAGAAC CTCAAGTAGG
18361 TGAAGAAAAC TGGCAAGAGT CTGAAGCCTT CTATGGAGGC AGAGCTCTTA AGAAAGACAC
18421 AAAAATGAAG CCCTGCTATG GCTCATTTGC AAGACCTACC AATGAAAAAG GCGGACAAGC
18481 TAAATTTAAG CCAGTGGAAG AGGGGCAGCA ACCTAAAGAT TATGACATAG ATTTGGCTTT
18541 CTTTGACACA CCTGGAGGCA CCATCACAGG AGGCACAGAC GAAGAATATA AAGCAGACAT
18601 TGTGTTGTAC ACTGAAAATG TCAACCTTGA AACCCCAGAC ACCCACGTGG TATACAAGCC
18661 AGGAAAAGAG GATGACAGTT CAGAAGTAAA TTTGACACAG CAGTCCATGC CCAACAGGCC
18721 TAACTACATT GGCTTCAGAG ACAACTTTGT GGGACTCATG TACTACAACA GTACTGGCAA
18781 CATGGGTGTG CTGGCTGGTC AGGCCTCTCA ATTGAATGCT GTGGTCGACT TGCAAGACAG
18841 AAACACCGAG CTGTCTTACC AGCTCTTGCT AGATTCTCTG GGTGACAGAA CCAGATACTT
18901 CAGCATGTGG AACTCTGCGG TGGATAGCTA TGATCCAGAT GTCAGGATCA TTGAAAATCA
18961 TGGTGTGGAA GATGAACTTC CAAACTATTG CTTCCCATTG AATGGCACTG GCACCAATTC
19021 AACATATCTT GGCGTAAAGG TGAAACCAGA TCAAGATGGT GATGTTGAAA GCGAGTGGGA
19081 TAAAGATGAT ACCATTGCAA GGCAGAATCA AATCGCCAAG GGCAACGTCT TTGCCATGGA
19141 GATCAACCTC CAGGCCAACC TGTGGAAGAG TTTTCTGTAC TCGAACGTGG CCTTGTACCT
19201 GCCCGACTCC TACAAGTACA CGCCGGCCAA TGTTACGCTG CCCGCCAACA CCAACACCTA
19261 CGAGTACATG AACGGCCGCG TGGTAGCCCC CTCGCTGGTG GACGCCTACA TCAACATAGG
19321 CGCCCGATGG TCGCTGGACC CCATGGACAA CGTCAACCCC TTCAACCACC ACCGCAATGC
19381 GGGCCTGCGC TACCGCTCCA TGCTTCTGGG CAACGGCCGC TACGTGCCCT TCCACATCCA
19441 AGTGCCCCAA AAGTTCTTTG CCATCAAGAA CCTGCTCCTG CTCCCGGGCT CCTACACCTA
19501 CGAGTGGAAC TTCCGCAAGG ATGTCAACAT GATCCTGCAG AGTTCCCTCG GCAACGACCT
19561 GCGCGTCGAC GGCGCCTCCG TCCGCTTCGA CAGCGTCAAC CTCTACGCCA CCTTCTTCCC
19621 CATGGCGCAC AACACCGCCT CCACCCTGGA AGCCATGCTG CGCAACGACA CCAACGACCA
19681 GTCCTTCAAC GACTACCTCT CGGCCGCCAA CATGCTCTAC CCCATCCCGG CCAAGGCCAC
19741 CAACGTGCCC ATCTCCATCC CCTCGCGCAA CTGGGCCGCT TTTCGCGGCT GGAGTTTCAC
19801 CCGTCTGAAA ACCAAGGAAA CTCCCTCCCT CGGCTCGGGT TTTGACCCCT ACTTTGTCTA
19861 CTCGGGCTCG ATCCCCTACC TTGACGGACC CTTTTACCTT AACCACACCT TCAAGAAAGT
19921 CTCCATCATG TTCGACTCCT CGGTCAGCTG GCCCGGCAAC GACCGGCTGC TCACGCCGAA
19981 CGAGTTCGAG ATCAAGCGCA GCGTCGACGG GGAAGGCTAC AACGTGGCCC AATGCAACAT
20041 GACCAAGGAC TGGTTCCTCG TCCAGATGCT CTCCCACTAC AACATCGGCT ACCAGGGCTT
20101 CCACGTGCCC GAGGGCTACA AGGACCGCAT GTACTCCTTC TTCCGCAACT TCCAGCCCAT
20161 GAGCAGGCAG GTGGTCGATG AGATCAACTA CAAGGACTAC AAGGCCGTCA CCCTGCCCTT
20221 CCAGCACAAC AACTCGGGCT TCACCGGCTA CCTTGCACCC ACCATGCGCC AAGGGCAGCC
20281 CTACCCCGCC AACTTCCCCT ACCCGCTCAT CGGCCAGACA GCCGTGCCAT CCGTCACCCA
20341 GAAAAGTCTC CTCTGCGACA GGGTCATGTG GCGCATCCCC TTCTCCAGCA ACTTCATGTC
20401 CATGGGCGCC TTCACCGACC TGGGTCAGAA CATGTTCTAC GCCAACTCGG CCCACGCGCT
20461 CGACATGACC TTCGAGGTGG ACCCCATGGA TGAGCCCACC GTCCTCTATC TTCTCTTCGA

20521 AGTGTTCGAC GTGGTCAGAG TGCACCAGCC GCACCGCGGC GTCATCGAGG CCGTCTACCT
20581 GCGCACGCCG TTCTCCGCCG GAAACGCCAC CACCTAAGCA TGAGCGGCTC CAGCGAAAGA
20641 GAGCTCGCGT CCATCGTGCG CGACCTGGGC TGCGGGCCTA CTTTTTGGGC ACCCACGACA
20701 CAGCGATTCC CGGGCTTTCT TGCCGGCGAC AAGCTGGCCT GCGCCATTGT CAACACGGCC
20761 GGCCGCGAGA CCGGAGGCGT GCACTGGCTC GCCTTCGGCT GGAACCCGCG CTCGCGCACC
20821 TGCTACATGT TCGACCCCTT TGGGTTCTCG GACCGCCGGC TCAAGCAGAT TTACAGCTTC
20881 GAGTACGAGG CCATGCTGCG CCGAAGCGCC GTGGCCTCTT CGCCCGACCG CTGTCTCAGC
20941 CTCGAACAGT CCACCCAGAC CGTGCAGGGG CCCGACTCCG CCGCCTGCGG ACTTTTCTGT
21001 TGCATGTTCT TGCATGCCTT CGTGCACTGG CCCGACCGAC CCATGGACGG GAACCCCACC
21061 ATGAACTTGC TGACGGGGGT GCCCAACGGC ATGCTACAAT CGCCACAGGT GCTGCCCACC
21121 CTCAGGCGCA ACCAGGAGGA GCTCTATCGC TTCCTCGCGC GCCACTCCCC TTACTTTCGC
21181 TCCCACCGCG CCGCCATCGA ACACGCCACC GCTTTTGACA AAATGAAACA ACTGCGTGTA
21241 TCTCAATAAA CAGCACTTTT ATTTTACATG CACTGGAGTA TATGCAAGTT ATTTAAAAGT
21301 CGAAGGGGTT CTCGCGCTCA TCGTTGTGCG CCGCGCTGGG GAGGGCCACG TTGCGGTACT
21361 GGTACTTGGG CTGCCACTTG AACTCGGGGA TCACCAGTTT GGGCACTGGG GTCTCGGGGA
21421 AGGTCTCGCT CCACATACGC CGGCTCATCT GCAGGGCGCC CAGCATGTCC GGGGCGGATA
21481 TCTTGAAATC GCAGTTGGGA CCGGTGCTCT GCGCGCGCGA GTTGCGGTAC ACGGGGTTGC
21541 AGCACTGGAA CACCATCAGA CTGGGGTACT TTACGCTGGC CAGCACGCTC TTGTCGCTGA
21601 TCTGATCCTT GTCCAGATCC TCGGCGTTGC TCACGCCGAA TGGGGTCATC TTGCACAGTT
21661 GGCGACCCAG GAATGGCACG CTCTGAGGCT TGTGGTTACA CTCGCAGTGC ACGGGCATCA
21721 GCATCATCCC CGCGCCGCGC TGCATATTCG GGTAGAGGCC TTGACAAAGG CCGTGATCTG
21781 CTTGAAAGCT TGTTGGGCCT TGGCCCCCTC GCTGAAAAAC AGGCCGCAGC TCTTCCCGCT
21841 GAACTGGTTA TTCCCGCACC CGGCATCCTG CACGCAGCAG CGCGCGTCAT GGCTGGTCAG
21901 TTGCACCACG CTTCTTCCCC AGCGGTTCTG GGTCACCTTG GCTTTGCTGG GTTGCTCCTT
21961 CAACGCGCGC TGCCCGTTCT CGCTGGTCAC ATCCATCTCC ACCACGTGGT CCTTGTGGAT
22021 CATCACCGTT CCATGCAGAC ACTTGAGCTG GCCTTCCACC TCGGTGCAGC CGTGATCCCA
22081 CAGGGCACTG CCGGTGCACT CCCAGTTCTT GTGCGCGATC CCGCTGTGGC TGAAGATGTA
22141 ACCTTGCAAG AGGCGACCCA TGATGGTGCT AAAGCTCTTC TGGGTGGTGA AGGTTAGTTG
22201 CAGACCGCGG GCCTCCTCGT TCATCCAGGT CTGGCACATC TTTTGGAAGA TCTCGGTCTG
22261 CTCGGGCATG AGCTTGTAAG CATCGCGCAG GCCGCTGTCG ACGCGGTAAC GTTCCATCAG
22321 CACGTTCATG GTATCCATGC CCTTTTCCCA GGACGAGACC AGAGGCAGAC TCAGGGGGTT
22381 GCGCACGTTC AGGACACCGG GGGTCKCGGG CTCGACGATA CGTTTTCCGT CCTTGCCTTC
22441 CTTCAACAGA ACCGGAGGCT GGCTGAATCC CACTCCCACA ATCACGGCAT CTTCCTGGGG
22501 CATCTCTTCG TCGGGGTCTA CCTTGGTCAC ATGCTTGGTC TTTCTGGCTT GCTTCTTTTT
22561 TGGAGGGCTG TCCACGGGGA CCACGTCCTC TCGGAAGACC CGGAGCCCAC CCGCTGATAC
22621 TTTCGGCGCT TGGTGGGCAG AGGAGGTGGC GGCGGCGAGG GGCTCCTCTC GTGCTCCGGC
22681 GGATAGCGCG CCGACCCGTG GCCCCGGGGC GGAGTGGCCT CTCGCTCCAT GAACCGGCGC
22741 ACGTCTGACT GCCGCCGGCC ATTGTTTCCT AGGGGAAGAT GGAGGAGCAG CCGCGTAAGC
22801 AGGAGCAGGA GGAGGACTTA ACCACCCACG AGCAACCCAA AATCGAGCAG GACCTGGGCT
22861 TCGAAGAGCC GGCTCGTCTA GAACCCCACA GGATGAACAG GAGCACGAGC AAGACGCAGG
22921 CCAGGAGGAG ACCGACGCTG GGCTCGAGCA TGGCTACCTG GGAGGAGAGG AGGATGTGCT
22981 GCTGAAACAC CTGCAGCGCC AGTCCCTCAT CCTCCGGGAC GCCCTGGCCG ACCGGAGCGA
23041 AACCCCCCTC AGCGTCGAGG AGCTGTGTCG GGCCTACGAG CTCAACCTCT TCTCGCCGCG
23101 CGTGCCCCCC AAACGCCAGC CCAACGGCAC CTGCGAGCCC AACCCGCGTC TCAACTTCTA
23161 TCCCGTCTTT GCGGTCCCCG AGGCCCTTGC CACCTATCAC ATCTTTTTCA AGAACCAAAA
23221 GATCCCCGTC TCCTGCCGCG CCAACCGCAC CCGCGCCGAC GCGCTCCTCG CTCTGGGGCC
23281 CGGCGCGCGC ATACCTGATA TTGCTTCCCT GGAAGAGTGC CCAAAATCTT CGAAGGGCTC
23341 GGTCGGGACG AGACGCGCGC GGCGAAACGC TCTGAAAGAA ACAGCAGAGG AAGAGGGTCA
23401 CACTAGCGCC CTGGTAGAGT TGGAAGGCGA CAACGCCAGG CTGGCCGTGC TCAAGCGCAG
23461 CGTTGAGCTC ACCCACTTCG CCTACCCCGC CGTCAACCTC CCGCCCAAGG TCATGCGTCG
23521 CATCATGGAT CAGCTAATCA TGCCCCACAT CGAGGCCCTC GATGAAAGTC AGGAGCAGCG
23581 CCCCGAGGAC ACCCGGCCCG TGGTCAGCGA TGAGCAGCTT GCGCGCTGGC TTGGTACCCG
23641 CGACCCCCAG GCCCTGGAGC AGCGGCGCAA GCTCATGCTG GCCGTGGTCC TGGTCACCCT
23701 CGAGCTCGAA TGCATGCGAC GCTTTTTCAG CGACCCCGAG ACCTGCGCAA GGTCGAGGAG
23761 ACCTGCACTA CACTTTTAGC ACGTTTCGTC AGGCAGGCAT GCAAGATCTC CAACGTGGAG
23821 CTGACCAACT GGTCTCCTGC CTGGGAATCC TGCACGAGAA CCGCCTGGGG CAGACAGTGC
23881 TCCACTCGAC CCTGAAGGGC GAGGCGCGGC GGGACTATGT CCGCGACTGC GTCTTTCTCT

23941 TTCTCTGCCA CACATGGCAA GCTGCCATGG GCGTGTGGCA GCAGTGTCTC GAGGACGAGA
24001 ACCTGAAGGA GCTGGACAAG CTTCTTGCTA GAAACCTCAA AAAGCTGTGG ACGGGCTTTG
24061 ACGAGCGCAC CGTCGCCTCG GACCTGGCCG AGATCGTCCT CCCCCGAGCG CCTGAGGCAG
24121 ACGCTGAAAG GCGGGCTGCC CGACTTCATG AGCCAGAGCA TGTTGCAAAA CTACCGCACT
24181 TTCATTCTCG AGCGATCTGG GATGCTGCCC GCCACCTGCA ACGCCTTCCC CTCCGACTTT
24241 GTCCCGCTGA GCTACCGCGA GTGTCCCCCG CCGCTGTGGA GCCACTGCTA CCTCTTGCAG
24301 CTGGCCAACT ACATCGCCTA CCACTCGGAT GTTATCGAGG ACGTGAGCGG CGAGGGGCTG
24361 CTAGAGTGCC ACTGCCGCTG CAACCTGTGC TCTCCGCACC GCTCCTGGTC TGCAACCCCC
24421 AGCTCCTGAG CGAGACCCAG GTCATCGGTA CCTTCGAGCT GCAAGGTCCG CAGGAGTCCA
24481 CCGCTCCGCT GAAACTCACG CCGGGGTTGT GGACTTCCGC GTACCTGCGC AAATTTGTAC
24541 CCGAGGACTA CCACGCCCAT GAGATAAAGT TCTTCGAGGA CCAATCGCGC CCGCAGCACG
24601 CGGATCTCAC GGCCTGCGTC ATCACCCAGG GCGCGATCCT CGCCCAATTG CACGCCATCC
24661 AAAAATCCCG CCAAGAGTTT CTTTTGAAAA AGGGTAGAGG GGTCTATCTG GACCCCCAGA
24721 CGGGCGAAGT GCTCAACCCG GGTCTCCCCC AGCATGCCGA AGAAGAACAG GAGCCGCTAG
24781 TGGAAGAGAT GGAAGAAGAA TGGGACAGCC AGCAGAAGAA GACGAATGGG AAGAAGAGAC
24841 AGAAGAAGAA GAATTGGAAA AGTGGAAGAA GAGCAGCACA GACACCGTCG CCGCACCATC
24901 CGCGCCGCAG CCCGGCGGTC ACGGATACAA CTCGCAGTCC GCCAAGCTCC TCGTAGATGG
24961 ATCGAGTGAA GGTGACGGTA AGCACGAGCG GCAGGGCTAC GAATCATGGA GGCCCACAAA
25021 GCGGGATCAT CGCCTGCTTG CAAGACTGCG GGGGGAACAT CGTTTCGCCC GCCGCTATCT
25081 GCTCTTCCAT CGCGGGGTGA ACATCCCCCG CAACGTGTTG CATTACTACC GTCACCTTCA
25141 CAGCTAAGAA AAAATCAGAG TAAGAGGAGT CGCCGGAGGA GGCNTGAGGA TCGCGGCGAA
25201 CGAGCCATTG ACCACCAGGG AGCTGAGGAA TCGGATCTTC CCCACTCTTT ATGCCATTTT
25261 TCAGCAGAGT CGAGGTCAGC AGCAAGAGCT CAAAGTAAAA AACCGGTCTC TGCGCTCGCT
25321 CACCCGCAGT TGCTTGTACC ACAAAAACGA AGATCAGCTG CAGCGCACTC TCGAAGACGC
25381 CGAGGCTCTG TTCCACAAGT ACTGCGCGCT CACTCTTAAA GACTAAGGCG CGCCCACCCG
25441 GAAAAAAGGC GGGAATTACC TCATCGCCAC CATGAGCAAG GAGATTCCCA CCCCTTACAT
25501 GTGGAGCTAT CAGCCCCAGA TGGGCCTGGC CGCGGGCGCC TCCCAGGACT ACTCCACCCG
25561 CATGAACTGG CTCAGTGCCG GCCCCTCGAT GATCTCACGG GTCAACGGGG TCCGTAACCA
25621 TCGAAACCAG ATATTGTTGG AGCAGGCGGC GGTCACCTCA ACGCCCAGGC AAAGCTCAAC
25681 CCGCGTAATT GGCCCTCCAC CCTGGTGTAT CAGGAAATCC CCGGGCCGAC TACCGTACTA
25741 CTTCCGCGTG ACGCACTGGC CGAAGTCCGC ATGACTAACT CAGGTGTCCA GCTGGCCGGC
25801 GGCGCTTCCC GGTGCCCGCT CCGCCCACAA TCGGGTATAA AAACCCTGGT GATACGAGGC
25861 AGAGGCACAC AGCTCAACGA CGAGTTGGTG AGCTCTTCAA TCGGTCTGCG ACCGGACGGA
25921 GTGTTCCAAC TAGCCGGAGC CGGGAGATCG TCCTTCACTC CCAACCAGGC TACCTGACCT
25981 TGCAGAGCAG CTCTTCGGAG CCTCGCTCCG GAGGCATCGG AACCCTCCAG TTTGTGGAGG
26041 AGTTTGTGCC CTCGGTCTAC TTCAACCCCT TCTCGGGATC GCCAGGCCTC TACCCGGACG
26101 AGTTCATACC GAACTTCGAC GCAGTGAGAG AAGCGGTGGA CGGCCACGAC TGAATGTCTT
26161 ATGGTGACTC GGCTGAGCTC GCTCGGTTGA GGCACCTAGA CCACTGCCGC CGCCTGCGCT
26221 GCTTCGCCCG GGAGAGCTGC GGACTTATCT ACTTTGAGTT TCCCGAGGAG CACCCCAACG
26281 GCCCTGCACA CGGAGTGCGG ATCACCGTAG AGGGCACCAC CGAGTCTCAC CTGGTTAGGT
26341 TCTTCACCCA GCAACCCTTC CTGGTCGAGC GGGACCGGGG AGGCACCACC TACACCGTCT
26401 ACTGCATCTG TCCAACCCCG AAGTTGCATG AGAATTTTTG TTGTACTCTG TGTGCTGAGT
26461 TTAATAAAAG CTAAACTCCT ACAATACTCT GGGATCCCGT GTCGTCGCAC TCGCAACAAG
26521 ACCTTCAACC TCACCAACCA GACTGAGGTA AAATTCAACT GCAGACCGGG GGACAAATAC
26581 ATCCTCTGGC TTTTTAAAAA CACTTCCTTC GCAGTCTCCA ACGCCTGCGC CAACGACGGT
26641 ATTGAAATAC CCAACAACCT TACCAGTGGA CTAACTTATA CTACCAGAAA GACTAAGCTA
26701 GTACTCTACA ATCCTTTTGT AGAGGGAACC TACCACTGCC AGAGCGGACC TTGCTTCCAC
26761 ACTTTCACTT TGGTGAACGT TACCGACAGC AGCACAGCCG CTACAGAAAC ATCTAACCTT
26821 CTTTTTGATA CTAACACTCC TAAAACCGGA GGTGAGCTCT GGGTTCCCTC TCTAACAGAG
26881 GGGGGTAAAC ATATTGAAGC GGTTGGGTAT TTGATTTTAG GGGTGGTCCT GGGTGGGTGC
26941 ATAGCGGTGC TGTATTACCT TCCTTGCTGG ATCGAAATCA AAATCTTTAT CTGCTGGGTC
27001 AGACATTGTT GGGAGGAACC ATGAAGGGGC TCTTGCTGAT TATCCTTTCC CTGGTGGGGG
27061 GTGTACTGTC ATGCCACGAA CAGCCACGAT GTAACATCAC CACAGGCAAT GAGAGGAGTG
27121 TGATATGCAC AGTAGTCATC AAATGCGAGC ATACATGCCC TCTCAACATC ACATTCAAAA
27181 ACCGTACCAT GGGAAATGCA TGGGTGGGCG ACTGGGAACC AGGAGATGAG CAGAACTACA
27241 CGGTCACTGT CCATGGTAGC AATGGAAATC ACACTTTTGG TTTCAAATTC ATTTTTGAAG
27301 TCATGTGTGA TATCACACTG CATGTGGCTA GACTTCATGG CTTGTGGCCC CCTACCAAGG

27361 ATAACATGGT TGGGTTTTCT TTGGCTTTTG TGATCATGGC CTGTGCAATG TCAGGTCTGC
27421 TGGTAGGGGC TTTAGTGTGG TTCCTAAAGC GCAAGCCTAG GTATGGAAAT GAGGAGAAGG
27481 AAAAATTGCT ATAAATCTTT TCTCTTCGCA GAACCATGAA TACAGTGATC CGTATCGTGC
27541 TGCTCTCTCT TCTTGTAACT TTTAGTCAGG CAGGATTCAT ACCATCAATG CTACATGGTG
27601 GGCTAATATA ACTTTAGTGG GACCTCAGAT ATTCCAGATC ACATGGTATG ATAGCACTGG
27661 ATTGCAATTT TGTGATGGAA GTACAGTTAA GAATCCACAG ATCAGACATA GTTGTAATGA
27721 TCAAAACTTA ACTCTGATTC ATGTGAACAA AACCCATGAA AGAACATACA TGGGCTATAA
27781 TAAGCAGAGT ACTCATAAAG AAGACTATAA AGTCACAGTT ATACCACCTC CTCCTGTTAC
27841 TGTAAAGCCA CAACCAGAGC CAGAATATGT GTATGTTAAT ATGGGAGAGA ACAAAACCTT
27901 AGTTGGGCCT CCAGGAATTC CAGTTAGTTG GTTTAATCAG GATGGTTTAC AATTTTGCAT
27961 TGGGGATAAA GTTTTTCATC CAGAATTCAA CCACACCTGT GACATGCAAA ATCTTACACT
28021 GTTGTTTATA AATCTTACAC ATGATGGAGC TTATCTTGGT TATAATCGCC AGGGAACTGA
28081 AAGAACTTGG TATGAGGTTG TAGTGTCAGA TGGTTTTCCA AAATCAGAAG AGATGAAGGT
28141 AGAAGACCAT AGTAAAGAAA CAGAACAAAA ACAGACTGGT CAAAAACAAA GTGACCATAA
28201 GCAGGGTGGG CAAAAAGAAA CAAGTCAAAA GAAAACTAAT GACAAACAAA AGCCATCGCG
28261 CAGGAGGCCA TCTAAACTAA AGCCAAACAC ACCTGACACA AAACTAATTA CAGTCACTAG
28321 TGGGTCAAAC GTAACTTTAG TTGGTCCAGA TGGAAAGGTC ACTTGGTATG ATGATGATTT
28381 AAAAAGACCA TGTGAGCCTG GGTATAAGTT AGGGTGTAAG TGTGACAATC AAAACCTAAC
28441 CCTAATCAAT GTAACTAAAC TTTATGAGGG AGTTTACTAT GGTACTAATG ACAGAGGCAA
28501 CAGCAAAAGA TACAGAGTAA AAGTAAACAC TACTAATTCT CAAAGTGTGA AAATTCAGCC
28561 GTACACCAGG CCTACTACTC CTGATCAGAA ACACAGATTT GAATTGCAAA TTGATTCTAA
28621 TCAAGACAAA ATTCCATCAA CTACTGTGGC AATCGTGGTG GGAGTGATCG CGGGCTTTGT
28681 AACTCTAATC ATTATTTTCA TATGCTACAT CTGCTGCCGC AAGCGTCCCA GGTCATACAA
28741 TCATATGGTA GACCCACTAC TCAGCTTCTC TTACTGAAAC TCAGTCACTC TCATTTCAGA
28801 ACCATGAAGG CTTTCACAGC TTGCGTTCTG ATTAGCATAG TCACACTTAG TTCAGCTGCA
28861 ATGATTAATG TTAATGTCAC TAGAGGTGGT AAAATTACAT TGAATGGGAC TTATCCACAA
28921 ACTACATGGA CAAGATATCA TAAAGATGGA TGGAAAAATA TTTGTGAATG GAATGTTACT
28981 GCATACAAAT GCTTCAATAA TGGAAGCATT ACTATTACTG CCACTGCCAA CATTACTTCT
29041 GGCACATACA AAGCTGAAAG CTATAAAAAT GAAATTAAAA AATTAACCTA TAAAAACAAC
29101 AAAACCACAT TTGAAGATTC TGGAAATTAT GAGCATCAAA AATTATCTTT TTATATGTTG
29161 ACAATAATTG AACTGCCTAC AACCAAGGCA CCCACCACAG TTAGTACAAC TACACAGTCA
29221 ACTGTTAAGA CCACTACTCA CACTACACAG CTAGACACCA CAGTGCAGAA TAATACTGTG
29281 TTGGTTAGGT ATTTGTTGAG GGAGGAAAGT ACTACTGAAC AGACAGAGGC TACCTCAAGT
29341 GCCTTTATCA GCACTGCAAA TTTAACTTCG CTTGCTTGGA CTAATGAAAC CGGAGTATCA
29401 TTGATGCATG GCCAGCCTTA CTCAGGTTTG GATATTCAAA TTACTTTTCT GGTTGTCTGT
29461 GGGATCTTTA TTCTTGTGGT TCTTCTGTAC TTTGTCTGCT GTAAAGCCAG AAAGAAATCT
29521 AGGAGGCCCA TCTACAGGCC AGTGATTGGG GAACCTCAGC CACTCCAAGT GGATGGAGGC
29581 TTAAGGAATC TTCTTTTCTC TTTTACAGTA TGGTGATCAG CCATGATTCC TAGTTCTTCC
29641 TATTTAACAT CCTCTTCTGT CTCTTCAACA TCTGTGCTGC CTTTGCGGCA GTTTCGCACG
29701 CCTCGCCCGA CTGTCTAGGG CCTTTCCCCA CCTACTCCTC TTTGCCCTGC TCACCTGCAC
29761 CTGCGTCTGC AGCATTGTCT GCCTGGTCAT CACCTTCCTG CAGCTCATCG ACTGGTGCTG
29821 CGCGCGCTAC AATTACTTCA TCATAGTCCC GAATACAGGG ACGAGAACGT AGCCAGAATT
29881 TTAAGGCTCA TATGACCATG CAGACTCTGC TCATACTGCT ATCGCTCTTA TCCCATGCCC
29941 TCGCTACTGC TGATTACTCT AAATGCAAAT TGGCGGACAT ATGGAATTTC TTAGACTGCT
30001 ATCAGGAGAA AATTGATATG CCCTCCTATT ACTTGGTGAT TGTGGGAATA GTTATGGTCT
30061 GCTCCTGCAC TTTCTTTGCC ATCATGATCT ACCCCTGTTT TGATCTTGGA TGGAACTCTG
30121 TTGAGGCATT CACATACACA CTAGAAAGCA GTTCACTAGC CTCCACGCCA CCACCCACAC
30181 CGCCTCCCCG CAGAAATCAG TTTCCCATGA TTCAGTACTT AGAAGAGCCC CCTCCCCGAC
30241 CCCCTTCCAC TGTTAGCTAC TTTCACATAA CCGGCGGCGA TGACTGACCA CCACCTGGAC
30301 CTCGAGATGG ACGGCCAGGC CTCCGAGCAG CGCATCCTGC AACTGCGCGT CCGTCAGCAG
30361 CAGGAGCGTG CCGCCAAGGA GCTCCTCGAT GCCATCAACA TCCACCAGTG CAAGAAGGGC
30421 ATCTTCTGCC TGGTCAAACA GGCAAAGATC ACCTACGAGC TCGTGTCCAA CGGCAAACAG
30481 CATCGCCTCA CCTATGAGAT GCCCCAGCAG AAGCAGAAGT TCACCTGCAT GGTGGGCGTC
30541 AACCCCATAG TCATCACCCA GCAGTCGGGC GAGACCAACG GCTGCATCCA CTGCTCCTGC
30601 GAAAGCCCCG AGTGTATCTA CTCCCTTCTC AAGACCCTTT GCGGACTCCG CGACCTCCTC
30661 CCCATGAACT GATGTTGATT AAAAACCAAA AAAAACAATC AGCCCCTTCC CCTATCCCAA
30721 ATTACTCGCA AAAATAAATC ATTGGAACTA ATCATTTAAT AAAGATCACT TACTTGAAAT

30781 CTGAAAGTAT GTCTCTGGTG TAGTTGTTCA GCAGCACCTC GGTACCCTCC TCCCAACTCT
30841 GGTACTCCAG TCTCCGGCGG GCGGCGAACT TTCTCCACAC CTTGAAAGGG ATGTCAAATT
30901 CCTGGTCCAC AATTTTCATT GTCTTCCCTC TCAGATGTCA AAGAGGCTCC GGGTGGAAGA
30961 TGACTTCAAC CCCGTCTACC CCTATGGCTA CGCGCGGAAT CAGAATATCC CCTTCCTCAC
31021 TCCCCCCTTT GTCTCCTCCG ATGGATTCAA AAACTTCCCC CCTGGGGTCC TGTCACTCAA
31081 ACTGGCTGAC CCAATCACCA TAGCCAATGG TGATGTCTCA CTCAAGGTGG GAGGGGACTT
31141 ACTTTGCAAG AAGGAAGTAT GACTGTAGAC CCTAAGGCTC CCTTGCAACT TGCAAACAAT
31201 AAAAAACTTG AGCTTGTTTA TGTTGATCCA TTTGAGGTTA GTGCCAATAA ACTTAGTTTA
31261 AAAGTAGGAC ATGGATTAAA AATATTAGAT GACAAAAGTG CTGGAGGGTT GAAAGATTTA
31321 ATTGGCAAAC TTGTGGTTTT AACAGGGGAA AGGAATAGGC ACTGAAAATT TGCAAAATAC
31381 AGATGGTAGC AGCAGAGGAA TTGGTATAAG TGTAAGAGCA AGAGAAGGGT TAACATTTGA
31441 CAATGATGGA TACTTGGTAG CATGGAACCC AAAGTATGAC ACGCGCACAC TTTGGACAAC
31501 ACCAGACACA TCTCCTAATT GCAGGATTGA TAAGGAGAAG ATTCAAAACT CACTTTGGTA
31561 CTTACAAAGT GTGGAAGTCA AATATTAGCT AATGTGTCTT TGATTGTGGT GTCAGGAAAA
31621 TATCAATACA TAGACCACGC TACAAATCCA ACTCTTAAAT CATTTAAAAT AAAACTTCTT
31681 TTTGATAATA AAGGTGTACT TCTCCCAAGT TCAAACCTTG ATTCCACATA TTGGAACTTT
31741 AGAAGTGACA ATTTAACTGT ATCTGAGGCA TATAAAAATG CAGTTGAATT TATGCCTAAT
31801 TTGGTAGCCT ACCCAAAACC TACCACTGGC TCTAAAAAAT ATGCAAGGGA TATAGTCTAT
31861 GGGAACATAT ATCTTGGAGG TTTGGCATAT CAGCCAGTTG TAATTAAGGT TACTTTTAAT
31921 GAAGAAGCAG ATAGTGCTTA CTCTATAACA TTTGAATTTG TATGGAATAA AGAATATGCC
31981 AGGGTTGAAT TTGAAACCAC TTCCTTTACC TTCTCCTATA TTGCCCAACA ATAAAAGACC
32041 AATAAACGTG TTTTTTATTT CAAATTTTAT GTATCTTTAT TGATTTTTAC ACCAGCGCGA
32101 GTAGTCAATC TCCCACCACC AGCCCATTTC ACAGTGTACA CGGTTCTCTC AGCACGGTGG
32161 CCTTAAATAA GGAAATGTTC TGATTATTGC GGGAACTGGA CTTGGGGTCT ATAATCCACA
32221 CAGTTTCCTG ACGAGCCAAA CGGGGATCGG TGATTGAAAT GAAGCCGTCC TCTGAAAAGT
32281 CATCCAAGCG GGCCTCACAG TCCAGGTCAC AGTCTGGTGG AACGAGAAGA ACGCACAGAT
32341 TCATACTCGG AAAACAGGAT GGGTCTGTGC CTCTCCATCA GCGCCCTCAG CAGTCTCTGC
32401 CGCCGGGGCT CGGTGCGGCT GCTGCAAATG GGATCGGGAT CACAAGTCTC TCTAACTATG
32461 ATCCCAACAG CCTTCAGCAT CAGTCTCCTG GTGCGTCGAG CACAGCACCG CATCCTGATC
32521 TCTGCCATGT TCTCACAGTA AGTGCAGCAC ATAATCACCA TGTTATTCAG CAGCCCATAA
32581 TTCAGGGTGC TCCAGCCAAA GCTCATGTTG GGGATGATGG AACCCACGTG ACCATCGTAC
32641 CAGATGCGGC AGTATATCAG GTGCCTGCCC CTCATGAACA CACTGCCCAT ATACATGATC
32701 TCTTTGGGCA TGTTTCTGTT TACAATCTGG CGGTACCAGG GGAAGCGCTG GTTGAACATG
32761 CACCCGTAAA TGACTCTCCT GAACCACACG GCCAGCAGGG TGCCTCCCGC CCGACACTGC
32821 AGGGAGCCAG GGGATGAACA GTGGCAATGC AGGATCCAGC GCTCGTACCC GCTCACCATC
32881 TGAGCTCTTA CCAAGTCCAG GGTAGCGGGG CACAGGCACA CTGACATACA TCTTTTTAAA
32941 ATTTTTATTT CCTCTGTGGT GAGGATCATA TCCCAGGGGA CTGGAAACTC TTGGAGCAGG
33001 GTAAAGCCAG CAGCACATGG TAATCCACGG ACAGAACTTA CATTATGATA ATCTGCATGA
33061 TCACAATCGG GCAACAGGGG ATGTTGATCA GTCAGTGAAG CCCTGGTTTC ATCATCAGAT
33121 CGTGGTAAAC GGGCCCTGCG ATATGGATGA TGGCGGAGCG AGCTGGATTG AATCTCGGTT
33181 TGCATTGTAG TGGATTCTCT TGCGTACCTT GTCGTACTTC TGCCAGCAGA AATGGGCCCT
33241 TGAACAGCAT ATACCCCTCC TGCGGCCGTC CTTTCGCTGC TGCCGCTCAG TCATCCAACT
33301 GAAGTACATC CATTCTCGAA GATTCTGGAG AAGTTCCTCT GCATCTGATG AAATAAAAAA
33361 CCCGTCCATG CGAATTCCCC TCATCACATC AGCCAGGACT CTGTAGGCCA TCCCCATCCA
33421 GTTAATGCTG CCTTGTCTAT CATTCAGAGG GGGCGGTGGC AGGATTGGAA GAACCATTTT
33481 TATTCCAAAC GGTCTCGAAG GACGATAAAG TGCAAGTCAC GCAGGTGACA GCGTTCCCCT
33541 CCGCTGTGCT GGTGGAAACA GACAGCCAGG TCAAAACCCA CTCTATTTTC AAGGTGCTCG
33601 ACCGTGGCTT CGAGCAGTGG CTCTACGCGT ACATCCAGCA TAAGAATCAC ATTAAAGGCT
33661 GGCCCTCCAT CGATTTCATC AATCATCAGG TTACATTCCT GCACCATCCC CAGGTAATTC
33721 TCATTTTTCC AGCCTTGGAT TATCTCTACA AATTGTTGGT GTAAATCCAC TCCGCACATG
33781 TTGAAAAGCT CCCACAGTGC CCCCTCCACT TTCATAATCA GGCAGACCTT CATAATAGAA
33841 ACAGATCCTG CTGCTCCACC ACCTGCAGCG TGTTCAAAAC AACAAGATTC AATAAGGTTC
33901 TGCCCTCCGC CCTGAGCTCG CGCCTCAATG TCAGCTGCAA AAAGTCACTT AAGTCCTGGG
33961 CCACTACAGC TGACAATTCA GAGCCAGGGC TAAGCGTGGG ACTGGCAAGC GTGAGGGAAA
34021 ACTTTAATGC TCCAAAGCTA GCACCCAAAA ACTGCATGCT GGAATAAGCT CTCTTTGTGT
34081 CTCCGGTGAT GCCTTCCAAA ATGTGAGTGA TAAAGCGTGG TAGTTTTTTC TTTAATCATT
34141 TGCGTAATAG AAAAGTCCTG TAAATAAGTC ACTAGGACCC CAGGGACCAC AATGTGGTAG

34201 CTTACACCGC GTCGCTGAAA GCATGGTTAG TAGAGATGAG AGTCTGAAAA ACAGAAAGCA
34261 TGCGCTAAAC TAAGGTGGCT ATTTTCACTG AAGGAAAAAT CACTCTTTCC AGCAGCAGGG
34321 TACCCACTGG GTGGCCCTTG CGGACATACA AAAATCGGTC CGTGTGATTA AAAAGCAGCA
34381 CAGTAAGTTC CTGTCTTCTT CCGGCAAAAA TCACATCGGA CTGGGTTAGT ATGTCCCTGG
34441 CATGGTAGTC ATTCAAGGCC ATAAATCTGC CCTGATATCC AGTAGGAACC AGCACACTCA
34501 CTTTTAGGTG AAGCAATACC ACCCCATGCG GAGGAATGTG GAAAGATTCA GGGCAAAAAA
34561 AATTATATCT ATTGCTAGCC CTTCCTGGAC GGGAGCAATC CTCCAGGACT ATCTATGAAA
34621 GCATACAGAG ATTCAGCCAT AGCTCAGCCC GCTTACCAGT AGACAAAGAG CACAGCAGTA
34681 CAAGCGCCAA CAGCAGCGAC TGACTACCCA CTGACTTAGC TCCCTATTTA AAGGCACCTT
34741 ACACTGACGT AATGACCAAA GGTCTAAAAA CCCCGCCAAA AAAACACACA CGCCCTGGGT
34801 GTTTTTGCGA AAACACTTCC GCGTTCTCAC TTCCTCGTAT CGATTTCGTG ACTTGACTTC
34861 CGGGTTCCCA CGTTACGTCA CTTTTGCCCT TACATGTAAC TTAGTCGTAG GGCGCCATCT
34921 TGCCCACGTC CAAAATGGCT TACATGTCCA GTTACGCCTC CGCGGCGACC GTTAGCCGTG
34981 CGTCGTGACG TCATTTGCAT CAACGTTTCT CGGCCAATCA GCAGTAGCCC CGCCCTAAAT
35041 TTAAAACCTC ATTTGCATAT TAACTTTTGT TTACTTTGTG GGGTATATTA TTGATGATG

ATGTCAAAGAGGCTCCGGGTGGAAGATGACTTCAACCCCGTCTACCCCTA
TGGCTACGCGCGGAATCAGAATATCCCCTTCCTCACTCCCCCCTTTGTCTC
CTCCGATGGATTCAAAAACTTCCCCCCTGGGGTCCTGTCACTCAAACTGGC
TGACCCAATCACCATAGCCAATGGTGATGTCTCACTCAAGGTGGGAGGGG
GACTTACTTTGCAAGAAGGAAGTCTGACTGTAGACCCTAAGGCTCCCTTG
CAACTTGCAAACAATAAAAAACTTGAGCTTGTTTATGTTGATCCATTTGAG
GTTAGTGCCAATAAACTTAGTTTAAAAGTAGGACATGGATTAAAAATATT
AGATGACAAAAGTGCTGGAGGGTTGAAAGATTTAATTGGCAAACTTGTGG
TTTTAACAGGGAAAGGAATAGGCACTGAAAATTTGCAAAATACAGATGGT
AGCAGCAGAGGAATTGGTATAAGTGTAAGAGCAAGAGAAGGGTTAACAT
TTGACAATGATGGATACTTGGTAGCATGGAACCCAAAGTATGACACGCGC
ACACTTTGGACAACACCAGACACATCTCCTAATTGCAGGATTGATAAGGA
GAAGGATTCAAAACTCACTTTGGTACTTACAAAGTGTGGAAGTCAAATAT
TAGCTAATGTGTCTTTGATTGTGGTGTCAGGAAAATATCAATACATAGACC
ACGCTACAAATCCAACTCTTAAATCATTTAAAATAAAACTTCTTTTTGATA
ATAAAGGTGTACTTCTCCCAAGTTCAAACCTTGATTCCACATATTGGAACT
TTAGAAGTGACAATTTAACTGTATCTGAGGCATATAAAAATGCAGTTGAA
TTTATGCCTAATTTGGTAGCCTACCCAAAACCTACCACTGGCTCTAAAAAA
TATGCAAGGGATATAGTCTATGGGAACATATATCTTGGAGGTTTGGCATA
TCAGCCAGTTGTAATTAAGGTTACTTTTAATGAAGAAGCAGATAGTGCTTA
CTCTATAACATTTGAATTTGTATGGAATAAAGAATATGCCAGGGGTTGAA
TTTGAAACCACTTCCTTTACCTTCTCCTATATTGCCCAACAATAA

SEQ ID NO:2

Penton17.Seq Length: 1554

1	ATGAGGCGTG	CGGTGGTGTC	TTCCTCTCCT	CCTCCCTCGT	ACGAGAGCGT
51	GATGGCGCAG	GCGACCCTGG	AGGTTCCGTT	TGTGCCTCCG	CGGTATATGG
101	CTCCTACGGA	GGGCAGAAAC	AGCATTCGTT	ACTCGGAGCT	GGCTCCGTTG
151	TACGACACCA	CTCGCGTGTA	CTTGGTGGAC	AACAAGTCGG	CGGACATCGC
201	TTCCCTGAAC	TATCAAAACG	ACCACAGCAA	CTTCCTGACC	ACGGTGGTGC
251	AGAACAACGA	TTTCACCCCC	GCCGAGGCTA	GCACGCAGAC	GATAAATTTT
301	GACGAGCGGT	CGCGGTGGGG	CGGTGATCTG	AAGACCATTC	TGCACACCAA
351	CATGCCCAAT	GTGAACGAGT	ACATGTTCAC	CAGCAAGTTT	AAGGCGCGGG
401	TGATGGTGGC	TAGAAAACAC	CCACAGGGGG	TAGAAGCAAC	AGATTTAAGC
451	AAGGATATCT	TAGAGTATGA	GTGGTTTGAG	TTTACCCTGC	CCGAGGGCAA
501	CTTTTCCGAG	ACCATGACCA	TAGACCTGAT	GAACAACGCC	ATCTTGGAAA
551	ACTACTTGCA	AGTGGGGCGG	CAAAATGGCG	TGCTGGAGAG	CGATATTGGA
601	GTCAAGTTTG	ACAGCAGAAA	TTTCAAGCTG	GGCTGGGACC	CTGTGACCAA
651	GCTGGTGATG	CCAGGGGTCT	ACACCTACGA	GGCCTTTCAC	CCGGACGTGG
701	TGCTGCTGCC	GGGCTGCGGG	GTGGACTTCA	CAGAGAGCCG	CCTGAGCAAC
751	CTCCTGGGCA	TTCGCAAGAA	GCAACCTTTC	CAAGAGGGCT	TCAGAATCAT
801	GTATGAGGAT	CTAGAAGGGG	GCAACATCCC	CGCCCTGCTG	GATGTGCCCA
851	AGTACTTGGA	AAGCAAGAAG	AAGTTAGAGG	AGGCATTGGA	GAATGCTGCT
901	AAAGCTAATG	GTCCTGCAAG	AGGAGACAGT	AGCGTCTCAA	GAGAGGTTGA
951	AAAGGCAGCT	GAAAAAGAAC	TTGTTATTGA	GCCCATCAAG	CAAGATGATA
1001	CCAAGAGAAG	TTACAACCTC	ATCGAGGGAA	CCATGGACAC	GCTGTACCGC
1051	AGCTGGTACC	TGTCCTATAC	CTACCGGGAC	CCTGAGAACG	GGGTGCAGTC
1101	GTGGACGCTG	CTCACCACCC	CGGACGTCAC	CTGCGGCGCG	GAGCAAGTCT
1151	ACTGGTCGCT	GCCGGACCTC	ATGCAAGACC	CCGTCACCTT	CCGTTCTACC
1201	CAGCAAGTCA	GCAACTACCC	CGTGGTCGGC	GCCGAGCTCA	TGCCCTTCCG
1251	CGCCAAGAGC	TTTTACAACG	ACCTCGCCGT	CTACTCCCAG	CTCATCCGCA
1301	GCTACACCTC	CCTCACCCAC	GTCTTCAACC	GCTTCCCCGA	CAACCAGATC

SEQ ID NO: 3

1351 CTCTGCCGTC CGCCCGCGCC CACCATCACC ACCGTCAGTG AAAACGTGCC

1401 TGCTCTCACA GATCACGGGA CGCTACCGCT GCGCAGCAGT ATCCGCGGAG

1451 TCCAGCGAGT GACCGTCACT GACGCCCGTC GCCGCACCTG TCCCTACGTC

1501 TACAAGGCCC TGGGCATAGT CGCGCCGCGT GTGCTTTCCA GTCGCACCTT

1551 CTAA

Claims

1. A chimeric adenoviral vector comprising nucleotide sequence of a first adenovirus, wherein at least one gene of said first adenovirus encoding a protein that facilitates binding of said vector to a target mammalian cell, or internalization thereof within said cell, is replaced by the corresponding gene from a second adenovirus belonging to subgroup D, said vector further comprising a transgene operably linked to a eucaryotic promoter to allow for expression therefrom in a mammalian cell.

2. A chimeric adenoviral vector according to Claim 1 wherein said second adenovirus is selected from the group consisting of Ad 9, Ad 15, Ad 17, Ad 19, Ad 20, Ad 22, Ad 26, Ad 27, Ad 28, Ad 30, and Ad 39.

3. A chimeric adenoviral vector according to Claim 1 wherein said first adenovirus is selected from the group consisting of Ad 2, Ad 5, and Ad 12.

4. A chimeric adenoviral vector according to Claim 1 wherein said replaced gene encodes Ad fiber.

5. A chimeric adenoviral vector according to Claim 1 wherein said replaced gene encodes Ad penton base.

6. A chimeric adenoviral vector according to Claim 1 wherein a first replaced gene encodes Ad fiber, and a second replaced gene encodes Ad penton base.

7. A chimeric adenoviral vector comprising nucleotide sequence of a first adenovirus, wherein a portion of a gene thereof encoding a protein that facilitates binding of said vector to a target mammalian cell, or internalization

thereof within said cell, is replaced by a portion of the corresponding gene from a second adenovirus belonging to subgroup D, said vector further comprising a transgene operably linked to a eucaryotic promoter to allow for expression therefrom in a mammalian cell.

8. A chimeric adenoviral vector according to Claim 7 wherein the encoding sequence that is replaced codes for a portion of Ad fiber.

9. A chimeric adenoviral vector according to Claim 7 wherein the encoding sequence that is replaced codes for a portion of Ad penton base.

10. A chimeric adenoviral vector according to Claim 9 wherein the encoding sequence that is replaced codes for an amino acid sequence that includes RGD.

11. A method of providing a biologically active protein to the airway epithelial cells of a patient comprising administering to said cells an adenoviral vector selected from the group consisting of:

(a) a chimeric adenoviral vector comprising nucleotide sequence of a first adenovirus, wherein at least one gene of said first adenovirus encodes a protein that facilitates binding of said vector to a target mammalian cell, or internalization thereof within said cell, is replaced by the corresponding gene from a second adenovirus belonging to subgroup D, said vector further comprising a transgene encoding said protein that is operably linked to a eucaryotic promoter to allow for expression therefrom in a mammalian cell; and

(b) a chimeric adenoviral vector comprising nucleotide sequence of a first adenovirus, wherein a portion of a gene thereof encoding a protein that facilitates binding of said vector to a target mammalian cell, or internalization thereof within said cell, is replaced by a portion of the

corresponding gene from a second adenovirus belonging to subgroup D, said vector further comprising a transgene encoding said protein that is operably linked to a eucaryotic promoter to allow for expression therefrom in a mammalian cell;

under conditions whereby the transgene encoding said protein is expressed, and phenotypic benefit is produced in said airway epithelial cells.

12. A method according to Claim 11 wherein said second adenovirus is Ad 17 and the nucleotide sequence thereof used in replacement of nucleotide sequence of said first adenovirus encodes a polypeptide selected from the group consisting of Ad 17 fiber, a fragment of Ad 17 fiber, Ad 17 hexon, a fragment of Ad 17 hexon, Ad penton base, and a fragment of Ad 17 penton base.

13. A method of providing a biologically active protein to the airway epithelial cells of a patient that comprises administering to said cells an adenoviral vector comprising elements of an Ad 17 genome, and a transgene encoding said protein that is operably linked to a eucaryotic promoter to allow for expression therefrom in a mammalian cell, under conditions whereby the transgene encoding said protein is expressed, and phenotypic benefit is produced in said airway epithelial cells.

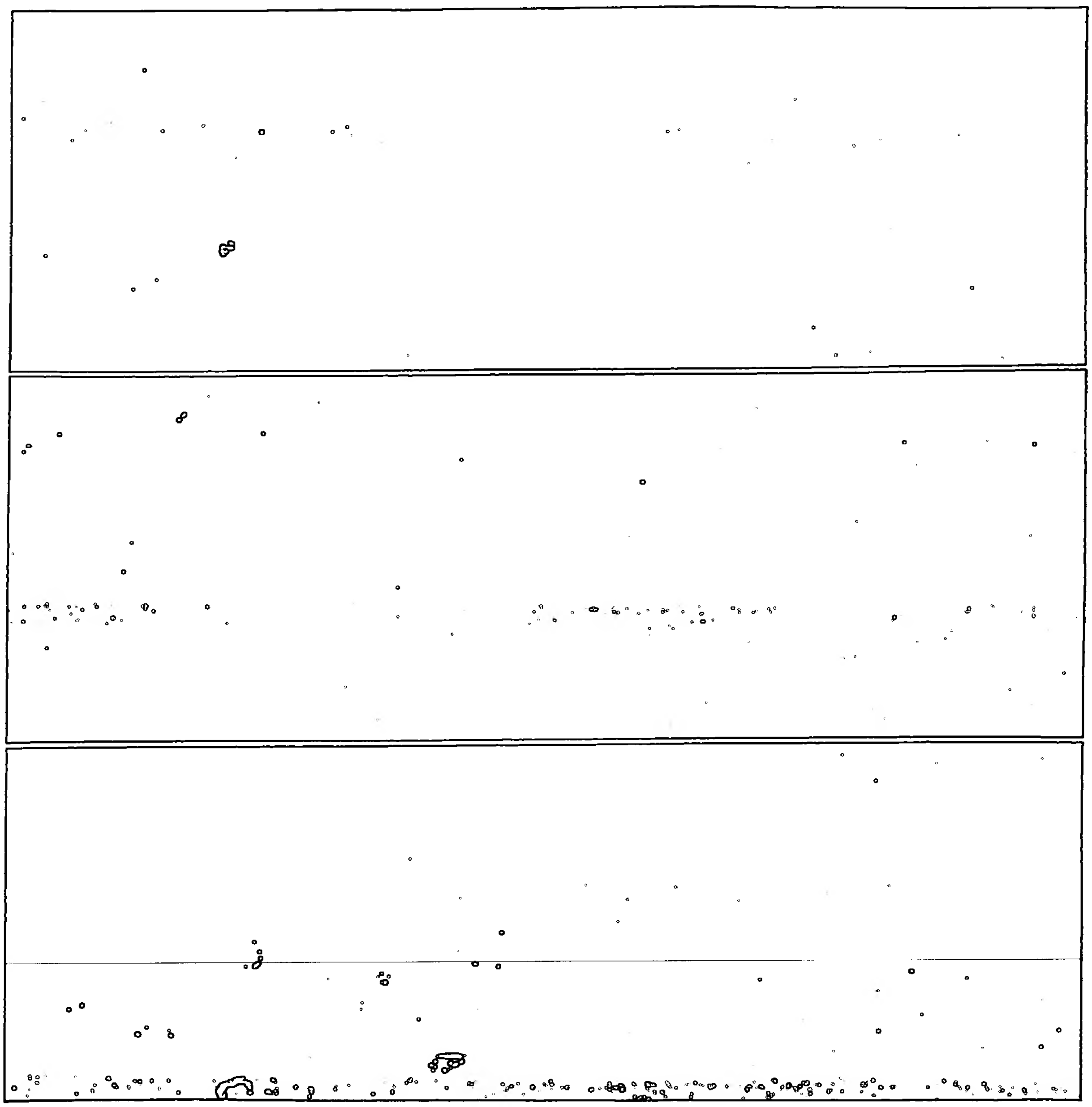


FIG 1 - original filed in PTO
is full color - see side folder

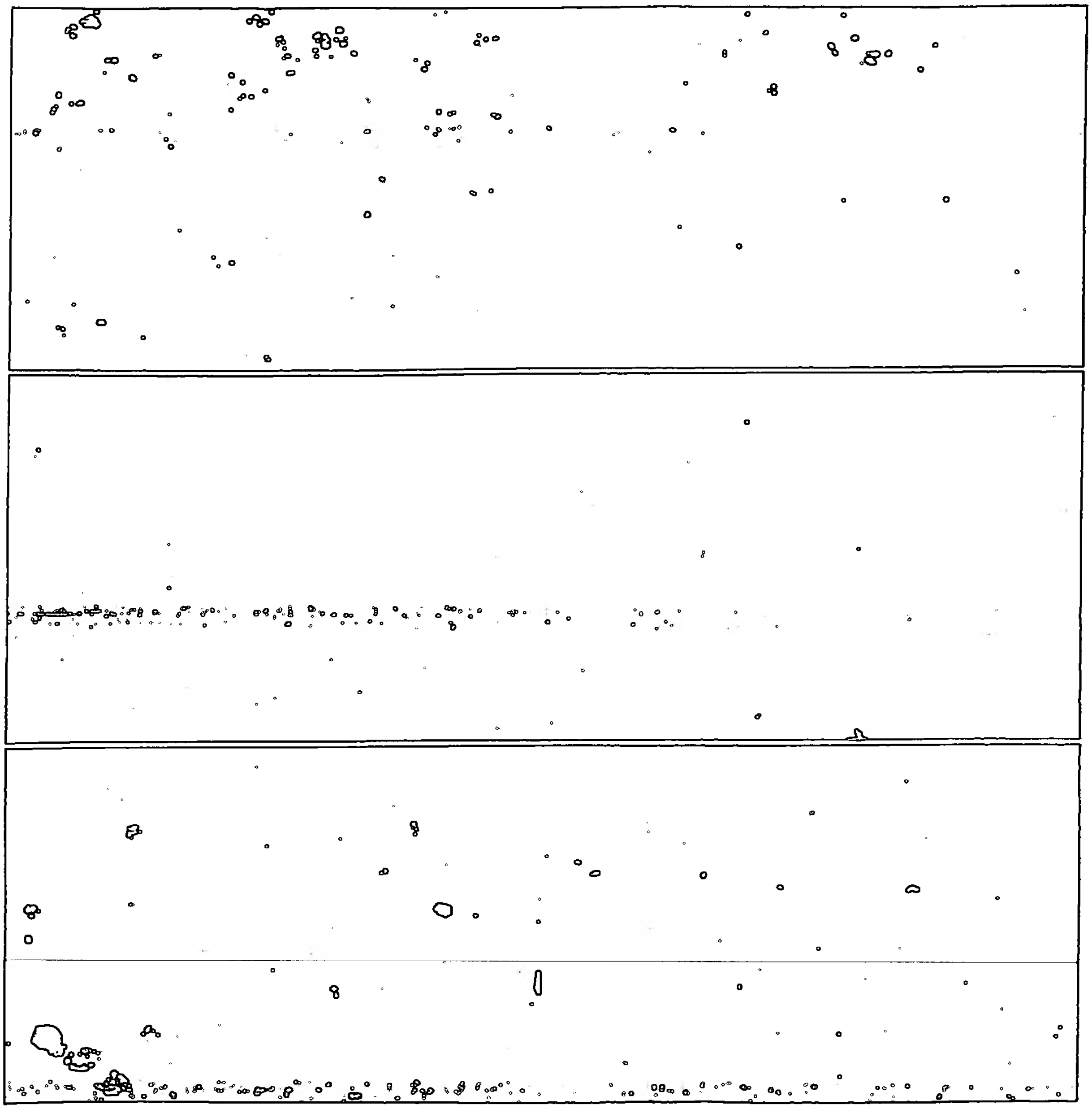


FIG 2 – original filed in PTO
is full color – see file folder

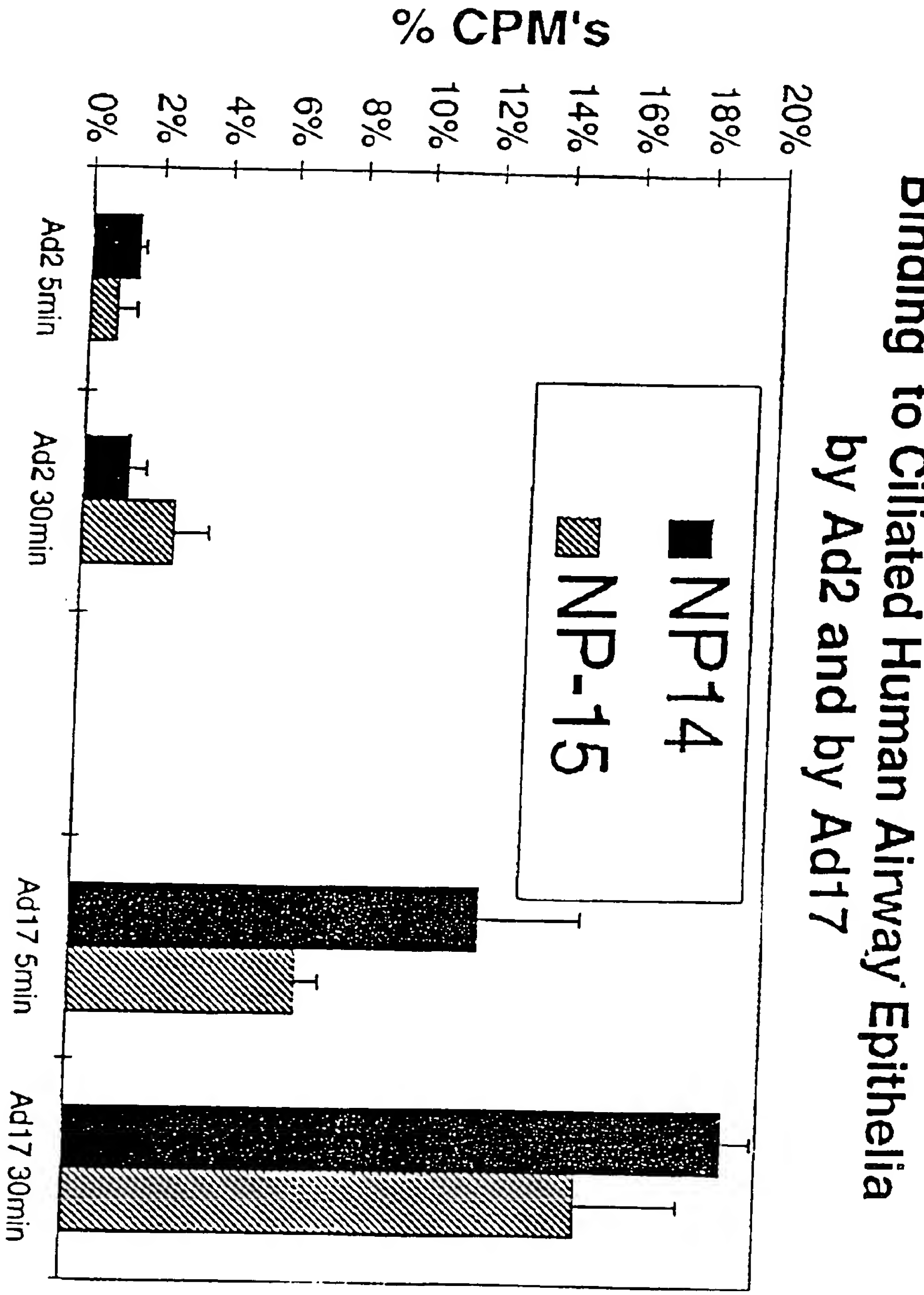

Binding to Ciliated Human Airway Epithelia by Ad2 and by Ad17
% CPM's
20%
18%
16%
14%
12%
10%
8%
6%
4%
2%
0%
NP14
NP-15
Ad2 5min
Ad2 30min
Ad17 5min
Ad17 30min

FIGURE 3

Chimeric Ad2/ßgal-2/ Ad17 vectors

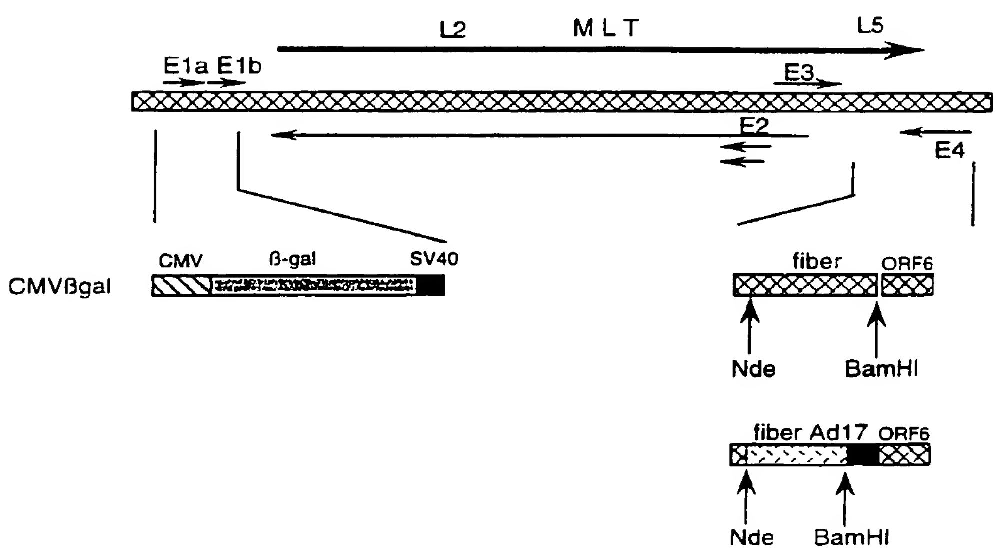


FIGURE 4

```
    MRRAVVSSSPPPSYESVMA............QATLEVPFVPPRYMAPTEGR  39    SEQ ID NO:4
  MQRAAMYEEGPPPSYESVVSAAPVAAALGSPFDAPLDPPFVPPRYLRPTGGR  52    SEQ ID NO:5

 40 NSIRYSELAPLYDTTRVYLVDNKSADIASLNYQNDHSNFLTTVVQNNDFT  89
 53 NSIRYSELAPLFDTTRVYLVDNKSTDVASLNYQNDHSNFLTTVIQNNDYS 102

 90 PAEASTQTINFDERSRWGGDLKTILHTNMPNVNEYMFTSKFKARVMVARK 139
103 PGEASTQTINLDDRSHWGGDLKTILHTNMPNVNEFMFTNKFKARVMVSRS 152

140 HPQGVEATDLSKDILEYEWFEFTLPEGNFSETMTIDLMNNAILENYLQVG 189
153 LTKDKQVE......LKYEWVEFTLPEGNYSETMTIDLMNNAIVEHYLKVG 196

190 RQNGVLESDIGVKFDSRNFKLGWDPVTKLVMPGVYTYEAFHPDVVLLPGC 239
197 RQNGVLESDIGVKFDTRNFRLGFDPVTGLVMPGVYTNEAFHPDIILLPGC 246

240 GVDFTESRLSNLLGIRKKQPFQEGFRIMYEDLEGGNIPALLDVPKYLES. 288   START
247 GVDFTHSRLSNLLGIRKRQPFQEGFRITYDDLEGGNIPALLDVDAYQASL 296

289 KKKLEEALENAAKANGPA.........................RGDSSVS 313
297 KDDTEQGGDGAGGGNNSGSGAEENSNAAAAAMQPVEDMNDHAIRGDTFAT 346

314 REVEKAAE...................KELVIEPIKQDDTKRSYNLIEG 343
347 RAEEKRAEAEAAAEAAAPAAQPEVEKPQKKPVIKPLTEDSKKRSYNLISN 396

344 TMD.TLYRSWYLSYTYRDPENGVCSWTLLTTPDVTCGAEQVYWSLPDLMQ 392
397 DSTFTQYRSWYLAYNYGDPQTGIFSWTLLCTPDVTCGSEQVYWSLPDMMQ 446

END
```

FIGURE 5A

```
393 DPVTFRSTQQVSNYPVVGAELMPFRAKSFYNDLAVYSQLIRSYTSLTHVF 442
    |||||||| |:||:|||||||:|.:.|||||| |||||||| :|||||||
447 DPVTFRSTSQISNFPVVGAELLPVHSKSFYNDQAVYSQLIRQFTSLTHVF 496

443 NRFPDNQILCRPPAPTITTVSENVPALTDHGTLPLRSSIRGVQRVTVTDA 492
    ||||:||||.||||||||||||||||||||||||||.|| ||||||:|||
497 NRFPENQILARPPAPTITTVSENVPALTDHGTLPLRNSIGGVQRVTITDA 546

493 RRRTCPYVYKALGIVAPRVLSSRTF 517
    |||||||||||||||.|||||||||
547 RRRTCPYVYKALGIVSPRVLSSRTF 571
```

FIGURE 5B

//

```
            1                                                  50
Penton5     ...MRRAAM. .....YEEGP PPSYESVVSA ..APVAAALG SPFDAPLDPP  ← SEQ ID NO: 6
Penton2     ...MQRAAM. .....YEEGP PPSYESVVSA ..APVAAALG SPFDAPLDPP  ← SEQ ID NO: 5
Penton3     ..MRRRAVLG GAV.VYPEGP PPSYESVM.. ......QQQA AMIQPPLEAP  ← SEQ ID NO: 7
Penton12    ...MRRAVEL QTV.AFPETP PPSYETVM.. .......... .....AAAPP  ← SEQ ID NO: 8
Penton40    ...MRRAVGV PPVMAYAEGP PPSYESVM.. ........ET ADLPATLQAL  ← SEQ ID NO: 9
Penton17    ...MRRAVV. ......SSSP PPSYESVM.. ......A... ...QATLEVP  ← SEQ ID NO: 4
Pentonf10   MWGLQPPTSI PPPPPPTELT PSTYPAMVNG YPPPAASAQS CSSSGGQSEL  ← SEQ ID NO: 10

            51                                                100
Penton5     FVP.PRYLRP TGGRNSIRYS ELAPLFDTTR VYLVDNKSTD VASLNYQNDH
Penton2     FVP.PRYLRP TGGRNSIRYS ELAPLFDTTR VYLVDNKSTD VASLNYQNDH
Penton3     FVP.PRYLAP TEGRNSIRYS DVSPLYDTTK LYLVDNKSAD IASLNYQNDH
Penton12    YVP.PRYLGP TEGRNSIRYS ELSPLYDTTR VYLVDNKSSD IASLNYQNDH
Penton40    HVP.PRYLGP TEGRNSIRYS ELAPLYDTTR VYLVDNKSAD IASLNYQNDH
Penton17    FVP.PRYMAP TEGRNSIRYS ELAPLYDTTR VYLVDNKSAD IASLNYQNDH
Pentonf10   YMPLQRVMAP TGGRNSIKYR DYTPCRNTTK LFYVDNKASD IDTYNKDANH

            101                                               150
Penton5     SNFLTTVIQN NDYSPGEAST QTINLDDRSH WGGDLKTILH TNMPNVNEFM
Penton2     SNFLTTVIQN NDYSPGEAST QTINLDDRSH WGGDLKTILH TNMPNVNEFM
Penton3     SNFLTTVVQN NDFTPTEAST QTINFDERSR WGGQLKTIMH TNMPNVNEYM
Penton12    SNFLTTVVQN NDYSPIEAGT QTINFDERSR WGGDLKTILH TNMPNVNDFM
Penton40    SNFQTTVVQN NDFTPTEAGT QTINFDDRSR WGGDLKTILR TNMPNINEFM
Penton17    SNFLTTVVQN NDFTPAEAST QTINFDERSR WGGDLKTILH TNMPNVNEYM
Pentonf10   SNFRTTVIHN QDLDADTAAT ESIQLDNRSC WGGDLKTAVR TNCPNVSSFF

            151                                               200
Penton5     FTNKFKARVM VSRL...... ...PTKD..N QVELKYEWVE FTLPEGNYSE
Penton2     FTNKFKARVM VSRS...... ...LTKD..K QVELKYEWVE FTLPEGNYSE
Penton3     FSNKFKARVM VSRKAPEGVT VNDTYDH..K EDILKYEWFE FILPEGNFSA
Penton12    FTTKFKARVM VARK...... ...TNNE..G QTILEYEWAE FVLPEGNYSE
Penton40    STNKFRARVM VEK....... ....VNR..K TNAPRYEWFE FTLPEGNYSE
Penton17    FTSKFKARVM VARKHPQGV. ..EATDL..S KDILEYEWFE FTLPEGNFSE
Pentonf10   QSNSVRVRMM WKRDPPTSTA PPSAVGSGYS VPGAQYKWYD LTVPEGNYAL

            201                                               250
Penton5     TMTIDLMNNA IVEHYLKVGR QNGVLESDIG VKFDTRNFRL GFDPVTGLVM
```

FIGURE 6A

```
  Penton2   TMTIDLMNNA IVEHYLKVGR QNGVLESDIG VKFDTRNFRL GFDPVTGLVM
  Penton3   TMTIDLMNNA IIDNYLEIGR QNGVLESDIG VKFDTRNFRL GWDPETKLIM
 Penton12   TMTIDLMNNA IIEHYLRVGR QHGVLESDIG VKFDTRNFRL GWDPETQLVT
 Penton40   TMTIDLMNNA IVDNYLAVGR QNGVLESDIG VKFDTRNFRL GWDPVTKLVM
 Penton17   TMTIDLMNNA ILENYLQVGR QNGVLESDIG VKFDSRNFKL GWDPVTKLVM
Pentonf10   CELIDLLNEG IVQLYLSEGR QNNVQKSDIG VKFDTRNFGL LRDPVTGLVT

            251                                               300
  Penton5   PGVYTNEAFH PDIILLPGCG VDFTHSRLSN LLGIRKRQPF QEGFRITYDD
  Penton2   PGVYTNEAFH PDIILLPGCG VDFTHSRLSN LLGIRKRQPF QEGFRITYDD
  Penton3   PGVYTYEAFH PDIVLLPGCG VDFTESRLSN LLGIRKRHPF QEGFKIMYED
 Penton12   PGVYTNEAFH PDIVLLPGCG VDFTESRLSN ILGIRKRQPF QEGFVIMYEH
 Penton40   PGVYTNEAFH PDIVLLPGCG VDFTQSRLNN LLGIRKRMPF QKGFQIMYED
 Penton17   PGVYTYEAFH PDVVLLPGCG VDFTESRLSN LLGIRKKQPF QEGFRIMYED
Pentonf10   PGTYVYKGYH PDIVLLPGCA IDFTYSRLSL LLGIGKREPY SKGFVITYED

            301                                               350
  Penton5   LEGGNIPALL DVDAYQASLK DDTEQGGGGA GGSNSSGSGA EENSNAAAAA
  Penton2   LEGGNIPALL DVDAYQASLK DDTEQGGDGA GGGNNSGSGA EENSNAAAAA
  Penton3   LEGGNIPALL DVTAYEESKK DTTTETTTLA VAEETSE... ..........
 Penton12   LEGGNIPALL DVKKYENSL. .......... ......Q... ..........
 Penton40   LEGGNIPALL DVEKYEASIK .......... .......... ..........
 Penton17   LEGGNIPALL DVPKYLESKK KLE......E ALENAAK... ..........
Pentonf10   LQGGDIPALL DLDSVDVNDA DGEVIELDNA A......... ..........

            351                                               400
  Penton5   MQPVEDMNDH AIRGDTFATR AEEKRAEAEA AAEAAAPAAQ PEVEKPQKKP
  Penton2   MQPVEDMNDH AIRGDTFATR AEEKRAEAEA AAEAAAPAAQ PEVEKPQKKP
  Penton3   .......DDD ITRGDTYITE KQKREAAAAE V......... ......KKEL
 Penton12   .......DQN TVRGDNFIA. .......... L......... ......NKAA
 Penton40   .......EAQ EIRGADFKPN PQ........ .......... ........DL
 Penton17   .......ANG PARGDSSVSR EVEKAA.... .......... ......EKEL
Pentonf10   .......... .......... .......... .......... ..........

            401                                               450
  Penton5   VIKPLTEDSK KRSYNLI... SNDSTFTQYR SWYLAYNYGD PQTGIRSWTL
  Penton2   VIKPLTEDSK KRSYNLI... SNDSTFTQYR SWYLAYNYGD PQTGIRSWTL
  Penton3   KIQPLEKDSK SRSYNVL... E.DKINTAYR SWYLSYNYGN PEKGIRSWTL
 Penton12   RIEPVETDPK GRSYNLL... P.DKKNTKYR SWYLAYNYGD PEKGVRSWTL
 Penton40   EIVPVEKDSK ERSYNLL... EGDKNNTAYR SWFLAYNYGD AEKGVKSWTL
 Penton17   VIEPIKQDDT KRSYNLI... E.GTMDTLYR SWYLSYTYRD PENGVQSWTL
Pentonf10   ...PLLHDSA GVSYNVIYDQ VTGKPVTAYR SWMLAYNVPN SQANQT..TL

            451                                               500
  Penton5   LCTPDVTCGS EQVYWSLPDM MQDPVTFRST RQISNFPVVG AELLPVHSKS
  Penton2   LCTPDVTCGS EQVYWSLPDM MQDPVTFRST SQISNFPVVG AELLPVHSKS
  Penton3   LTTSDVTCGA EQVYWSLPDM MQDPVTFRST RQVNNYPVVG AELMPVFSKS
 Penton12   LTTPDVTGGS EQVYWSLPDM MQDPVTFRSS RQVSNYPVVA AELLPVHAKS
 Penton40   LTTTDVTCGS QQVYWSLPDM MQDPVTFRPS TQVSNYPVVG VELLPVHAKS
 Penton17   LTTPDVTCGA EQVYWSLPDL MQDPVTFRST QQVSNYPVVG AELMPFRAKS
Pentonf10   LTVPDMAGGI GAMYTSLPDT FIAPTGFKED NTTNLCPVVG MNLFPTYNKI

            501                                               550
  Penton5   FYNDQAVYSQ LIRQFT.SLT HVFNRFPENQ ILARPPAPTI TTVSENVPAL
  Penton2   FYNDQAVYSQ LIRQFT.SLT HVFNRFPENQ ILARPPAPTI TTVSENVPAL
  Penton3   FYNEQAVYSQ QLRQAT.SLT HVFNRFPENQ ILIRPPAPTI TTVSENVPAL
 Penton12   FYNEQAVYSQ LIRQST.ALT RVFNRFPENQ ILVRPPAATI TTVSENVPAL
```

FIGURE 6B

```
  Penton40  FYNEQAVYSQ LIRQST.ALT HIFNRFPENQ ILVRPPAPTI TTVSENVPAL
  Penton17  FYNDLAVYSQ LIRSYT.SLT HVFNRFPDNQ ILCRPPAPTI TTVSENVPAL
 Pentonf10  YYQAASTYVQ RLENSCQSAT AAFNRFPENE ILKQAPPMNV SSVCDNQPAV

            551                                              600
   Penton5  TDHGTLPLRN SIGGVQRVTI TDARRRTCPY VYKALGIVSP RVLSSRTF*.
   Penton2  TDHGTLPLRN SIGGVQRVTI TDARRRTCPY VYKALGIVSP RVLSSRTF*.
   Penton3  TDHGTLPLRS SIRGVQRVTV TDARRRTCPY VYKALGIVAP RVLSSRTF*.
  Penton12  TDHGTLPLRS SISGVQRVTI TDARRRTCPY VYKALGIVSP RVLSSRTF*.
  Penton40  TDHGTLPLRS SISGVQRVTI TDARRRTCPY VHKALGIVAP KVLSSRTF*.
  Penton17  TDHGTLPLRS SIRGVQRVTV TDARRRTCPY VYKALGIVAP RVLSSRTF*.
 Pentonf10  VQQGVLPVKS SLPGLQRVLI TDDQRRPIPY VYKSIATVQP TVLSSATLQ*
```

FIGURE 6C

Fiber17.Pep x Fiber2.Pep

```
  1 MSKRLRVEDDFNPVYPYGYARN.QNIPFLTPPFVSSDGFKNFPPGVLSLK  49 ← SEQ ID:11
    |.:  . ||.|||||||: . . ..:|||||||||.:||.: |||||||:
  1 MKRARPSEDTFNPVYPYDTETGPPTVPFLTPPFVSPNGFQESPPGVLSLR  50 ← SEQ ID NO:12

 50 LADPITIANGDVSLKVGGGLTLQE..........................  73
    :.:|:...:| :.||:|:||||:.
 51 VSEPLDTSHGMLALKMGSGLTLDKAGNLTSQNVTTVTQPLKKTKSNISLD 100

 74 .........GSLTVDPKAPLQLA..............NNKKLELVYVDPF 100
             |.|||...||| :.              .:.||.:.  :|:
101 TSAPLTITSGALTVATTAPLIVTSGALSVQSQAPLTVQDSKLSIATKGPI 150

101 EVSANKLSLKVGHGLK.............................ILDDK 121
    .||..||.|..: .|.                             |. :.
151 TVSDGKLALQTSAPLSGSDSDTLTVTASPPLTTATGSLGINMEDPIYVNN 200

122 SAGGLK........DLIGKLVVLTGKGIGTE................... 144
    :  |:|        :  :.|.|:||.|:..|
201 GKIGIKISGPLQVAQNSDTLTVVTGPGVTVEQNSLRTKVAGAIGYDSSNN 250

                         .
                         .
                         .

145 ...........................NLQNTD...GSSRGIGISVRARE 164
                               |::||.   ..::|:::.....|
301 YNRGLYLFNASNNTKKLEVSIKKSSGLNFDNTAIAINAGKGLEFDTNTSE 350

165 .............................GLTFDNDGYLVAWNPKYDTRT 185
                                 ||.|||.| :.  |.. |. |
351 SPDINPIKTKIGSGIDYNENGAMITKLGAGLSFDNSGAITIGNKNDDKLT 400

186 LWTTPDTSPNCRIDKEKDSKLTLVLTKCGSQILANVSLIVVSGKYQYIDH 235
    ||||||.||||||..:.|:|:||||||||||:||.|. :.|||..
401 LWTTPDPSPNCRIHSDNDCKFTLVLTKCGSQVLATVAALAVSGDLS.... 446
```

FIGURE 7A

```
236 ATNPTLKSFKIKLLFDNKGVLLPSSNLDSTYWNFRSDNLTVSEAYKNAVE 285
    . ..|: |..| | ||..|||:..|.|.. |||||.:| | .::|.|||:
447 SMTGTVASVSIFLRFDQNGVLMENSSLKKHYWNFRNGNSTNANPYTNAVG 496
                .         .         .         .         .
286 FMPNLVAYPKPTTGSKKYARDIVYGNIYLGGLAYQPVVIKVTFNEEAD.. 333
    |||||:||||. ...    |:: : :.:|| |    .|::::.:|:|:...:
497 FMPNLLAYPKTQSQT...AKNNIVSQVYLHGDKTKPMILTITLNGTSEST 543
                .         .         .
334 .....SAYSITFEFVWNKE.YARVEFETTSFTFSYIAQQ 366
         |.||:.|.: |:.: |.  .|.|.|:||||||||:
544 ETSEVSTYSMSFTWSWESGKYTTETFATNSYTFSYIAQE 582
```

FIGURE 7B

```
            1                                                  50
8fiber      MTKRLRA... .....EDDFN PVYPYGYARN Q.NIPFLTPP FVSSNGFQNF - SEQ ID NO:13
9fiber      MSKRLRV... .....EDDFN PVYPYGYARN Q.NIPFLTPP FVSSDGFQNF - SEQ ID NO:14
15fiber     MSKRLRV... .....EDDFN PVYPYGYARN Q.NIPFLTPP FVSSDGFQNF - SEQ ID NO:15
17fiber     MSKRLRV... .....EDDFN PVYPYGYARN Q.NIPFLTPP FVSSDGFKNF - SEQ ID NO:11
2fiber      .MKRARP... ....SEDTFN PVYPYDTETG PPTVPFLTPP FVSPNGFQES - SEQ ID NO:12
5fiber      .MKRARP... ....SEDTFN PVYPYDTETG PPTVPFLTPP FVSPNGFQES - SEQ ID NO:16
4fiber      MSKSARG... ....WSDGFD PVYPYDADND RP.CPSSTLP SFSSDGFQEK - SEQ ID NO:17
40-1fiber   .MKRTRIE.. ......DDFN PVYPYD.TSS TPSIPYVAPP FVSSDGLQEN - SEQ ID NO:18
41fiber     .MKRTRIE.. ......DDFN PVYPYD.TFS TPSIPYVAPP FVSSDGLQEK - SEQ ID NO:19
40-2fiber   .MKRARFE.. ......DDFN PVYPYE.HYN PLDIPFITPP FASSNGLQEK - SEQ ID NO:20
12fiber     .MKRSRTQYA EETEENDDFN PVYPFD.PFD TSDVPFVTPP FTSSNGLQEK - SEQ ID NO:21
3fiber      MAKRARL... .....STSFN PVYPYEDESS SQH.PFINPG FISPDGFTQS - SEQ ID NO:22

            51                                                 100
8fiber      PPGVLSLKLA DPITIN.NQN VSLKVGGGLT LQEET..... ..........
9fiber      PPGVLSLKLA DPIAIV.NGN VSLKVGGGLT LQDGT..... ..........
15fiber     PPGVLSLKLA DPIAIA.NGN VSLKMGGGLT LQEGT..... ..........
17fiber     PPGVLSLKLA DPITIA.NGD VSLKVGGGLT LQE....... ..........
2fiber      PPGVLSLRVS EPLDTS.HGM LALKMGSGLT LDKAGNLTSQ NVTTVTQPLK
5fiber      PPGVLSLRLS EPLVTS.NGM LALKMGNGLS LDEAGNLTSQ NVTTVSPPLK
4fiber      PLGVLSLGPG RPCHTK.NGE ITLKLGEGVD LDDSGKLIAN TVNKAIAPL.
40-1fiber   PPGVLALKYT DPITTNAKHE LTLKLGSNIT LQ.NGLLSA. ..........
41fiber     PPGVLALKYT DPITTNAKHE LTLKLGSNIT LE.NGLLSA. ..........
40-2fiber   PPGVLSLKYT DPLTTK.NGA LTLKLGTGLN IDKNGDLSSD ASVEVSAPIT
12fiber     PPGVLALNYK DPIVTE.NGT LTLKLGDGIK LNAQGQLTAS NNINVLEPLT
3fiber      PNGVLSLKCV NPLTTA.SGS LQLKVGSGLT VD........ ..........

            101                                                150
8fiber      .......... .......... .......... .......... ..........
9fiber      .......... .......... .......... .......... ..........
15fiber     .......... .......... .......... .......... ..........
17fiber     .......... .......... .......... .......... ..........
```

FIGURE 8A

```
     2fiber  KTKSNISLD. S   LTITSGA LTVATTAPLI VTSGALS  . QAPLTVQDSK
     5fiber  KTKSNINLEI SAPLTVTSEA LTVAAAAPLM VAGNTLTMQS QAPLTVHDSK
     4fiber  .......... .......... .......... ....SFFQQH HFPL......
  40-1fiber  .......... .......... .......... .......... ..........
    41fiber  .......... .......... .......... .......... ..........
  40-2fiber  KTNKIVGLNY TKPLALQNNA LTLSYNAPFN VVNNNLALNM SQPVTI....
    12fiber  NTSQGLKLSW SAPLAVKASA LTLNTRAPLT TTDESLALIT APPITVESSR
     3fiber  .......... .......... .......... .......... ..........

             151                                                200
     8fiber  .......... .......... .......... .......... ..........
     9fiber  .......... .......... .......... .......... ..........
    15fiber  .......... .......... .......... .......... ..........
    17fiber  .......... .......... .......... .......... ..........
     2fiber  .......... .......... .......... .......... .......LSI
     5fiber  .......... .......... .......... .......... .......LSI
     4fiber  .......... .......... .......... .......... ..........
  40-1fiber  .......... .......... .......... .......... ..........
    41fiber  .......... .......... .......... .......... ..........
  40-2fiber  .......... .NANNELSLL IDAPLNADTG TLRLRSDAPL GLVDK.TLKV
    12fiber  LGLATIAPLS LDGGGNLGLN LSAPLDVSNN NLHLTTETPL VVNSSGALSV
     3fiber  .......... .......... .......... .......... ..........

             201                                                250
     8fiber  .......... .......... ......GKLT VNTEPPLH.. ..........
     9fiber  .......... .......... ......GKLT VNADPPLQ.. ..........
    15fiber  .......... .......... ......GNLT VNTEPPLQ.. ..........
    17fiber  .......... .......... ......GSLT VDPKAPLQ.. ..........
     2fiber  ATKGPITVSD GKLALQTSAP LSGSDSDTLT VTASPPLTTA TGSLGINMED
     5fiber  ATQGPLTVSE GKLALQTSGP LTTTDSSTLT ITASPPLTTA TGSLGIDLKE
     4fiber  .......... ......TWIP LYTPKMENYP YKFLPPLSIL KSTI......
  40-1fiber  .......... ...TVPT... .......... ..VSPPLTNS NNSLGLATSA
    41fiber  .......... ...TVPT... .......... ..VSPPLTNS NNSLGLATSA
  40-2fiber  LFSSPLYLDN NFLTLAIERP LALSSNRAVA LKYSPPLKIE NENLTLSTGG
    12fiber  ATADPISVRN NALTLPTADP LMVSSD.GLG ISVTSPITVI NGSLALSTTA
     3fiber  .......... .......... .......... .......... ..........

             251                                                300
     8fiber  ..LTNN.KLG IALDAPFDVI D..NKLTLLA GHGLSII.TK ETSTLPGLVN
     9fiber  ..LTNN.KLG IALDAPFDVI D..NKLTLLA GHGLSII.TK ETSTLPGLRN
    15fiber  ..LTNN.RIG IALDAPFDVI G..GKLTLLA GHGLSII.TE ETSPLPGLVN
    17fiber  ..LANNKKLE LVYVDPFEVS A..NKLSLKV GHGLKILDDK SAGGLKDLIG
     2fiber  PIYVNNGKIG IKISGPLQVA QNSDTLTVVT GPGVTVEQNS LRTKVAGAIG
     5fiber  PIYTQNGKLG LKYGAPLHVT DDLNTLTVAT GPGVTINNTS LQTKVTGALG
     4fiber  .......... .......... ..LNTLVSAF GSGLGLSGSA LAVQLASPLT
  40-1fiber  PIAVSANSLT LATAAPLTVS N..NQLSINT GRGLVITNNA VAVNPTGALG
    41fiber  PIAVSANSLT LATAAPLTVS N..NQLSINA GRGLVITNNA LTVNPTGALG
  40-2fiber  PFTVSGGNLN LATSAPLSVQ N..NSLSLGV NPPFLITDSG LAMDLGDGLA
    12fiber  PLNSTGSTLS LSVANPLTIS Q..DTLTVST GNGLQVSGSQ LVTRIGDGLT
     3fiber  ...TTDGSLE ENIKVNTPLT KSNHSINLPI GNGLQIEQNK LCS.......

             301                                                350
     8fiber  .......... .......... .......... .......... ..........
     9fiber  .......... .......... .......... .......... ..........
    15fiber  .......... .......... .......... .......... ..........
    17fiber  .......... .......... .......... .......... ..........
     2fiber  YDSSNNMEIK TGGGMRIN.. NNLLILDVDY PFDAQTKLRL KLGQGPLYIN
```

FIGURE 8B

```
    5fiber  FDSQGNMQLN [illegible]LRIDSQ NRRLILDVSY PFDAQNQL   RLGQGPLFIN
    4fiber  FDDKG..... .......... .......... .......... ..........
 40-1fiber  FNNTGALQLN AAGGMRVDGA N..LILHVAY PFEAINQLTL R.........
   41fiber  FNNTGALQLN AAGGMRVDGA N..LILHVAY PFEAINQLTL R.........
 40-2fiber  LGG.SKLIIN LGPGLQMSNG A..ITL.... ALDAALPL.. .........Q
   12fiber  FDN.GVMKVN VAGGMRTSGG R..IILDVNY PFDASNNLSL RRGLGLIYNQ
    3fiber  .......... .......... .......... .......... ..........

            351                                               400
    8fiber  .......... .......... .......... .......... TLVVLTGKGI
    9fiber  .......... .......... .......... .......... TLVVLTGKGI
   15fiber  .......... .......... .......... .......... TLVVLTGKGL
   17fiber  .......... .......... .......... .......... KLVVLTGKGI
    2fiber  ASHNLDINYN RGLYLFNASN NTKKLEVSIK KSSGLNFDNT AIAINAGKGL
    5fiber  SAHNLDINYN KGLYLFTASN NSKKLEVNLS TAKGLMFDAT AIAINAGDGL
    4fiber  ...NIKITLN RGLHVTTGDA ....IESNIS WAKGIKFEDG AIATNIGKGS
 40-1fiber  .......... .......... .......... .......... ..........
   41fiber  .......... .......... .......... .......... ..........
 40-2fiber  YKNN...... .......... .......... .......... ..QLQLRIGS
   12fiber  STNW...... .......... .......... .......... ..NLTTDIST
    3fiber  .......... .......... .......... .......... ..........

            401                                               450
    8fiber  GTDLSNNGG. ...NICVRVG E......... .....GGGLS FNDNGDLVAF
    9fiber  GTESTDNGG. ...TVCVRVG E......... .....GGGLS FNNDGDLVAF
   15fiber  GTDTTDNGG. ...SIRVRVG E......... .....GGGLS FNEAGDLVAF
   17fiber  GTENLQNTDG SSRGIGISVR A......... .....REGLT FDNDGYLVAW
    2fiber  EFDTNTSESP DINPIKTKIG SGIDYNENGA MITKLGAGLS FDNSGAITIG
    5fiber  EFG..SPNAP NTNPLKTKIG HGLEFDSNKA MVPKLGTGLS FDSTGAITVG
    4fiber  RFGTSSTETG VNNAYPIQV. .......... ...KLGSGLS FDSTGAIMAG
 40-1fiber  .......... ........LE NGLEVTNGGK LNVKLGSGLQ FDNNGRITIS
   41fiber  .......... ........LE NGLEVTSGGK LNVKLGSGLQ FDSNGRIAIS
 40-2fiber  ASALIMSGVT QTLNVNANTS KGLAIENNS. LVVKLGNGLR FDSWGSIAVS
   12fiber  EKGLMFSGN. ...QIALNAG QGLTFNNGQ. LRVKLGAGLI FDSNNNIALG
    3fiber  .......... .......... .......... ...KLGNGLT FDSSNSIALK

            451                                               500
    8fiber  NKKEDK.... .RTLWTTPDT SPNCRID... QDKDSKLSLV LTKCGSQILA
    9fiber  NKKEDK.... .RTLWTTPDT SPNCKID... QDKDSKLTLV LTKCGSQILA
   15fiber  NKKEDM.... .RTLWTTPDP SPNCKII... EDKDSKLTLI LTKCGSQILG
   17fiber  NPKYDT.... .RTLWTTPDT SPNCRID... KEKDSKLTLV LTKCGSQILA
    2fiber  NKNDDK.... .LTLWTTPDP SPNCRIH... SDNDCKFTLV LTKCGSQVLA
    5fiber  NKNNDK.... .LTLWTTPAP SPNCRLN... AEKDAKLTLV LTKCGSQILA
    4fiber  NKDYDK.... .LTLWTTPDP SPNCQIL... AENDAKLTLC LTMCDSQILA
 40-1fiber  NRIQTRSVTS LTTIWSIS.P TPNCSIY... ETQDANLFLC LTKNGAHVLG
   41fiber  NSNRTRSVPS LTTIWSIS.P TPNCSIY... ETQDANLFLC LTKNGAHVLG
 40-2fiber  PTTTT...P. .TTLWTTADP SPNATFY... ESLDAKVWLV LVKCNGMVNG
   12fiber  SSSNTPYDP. .LTLWTTPDP PPNCSLI... QELDAKLTLC LTKNGSIVNG
    3fiber  NN........ ..TLWTGPKP EANCIIEYGK QNPDSKLTLI LVKNGGIVNG

            501                                               550
    8fiber  NVSLIVVAGR YKIINNNTNP ..ALKGFTIK LLFDKNGVLM ESSN......
    9fiber  NVSLIVVDGK YKIINNNTQP ..ALKGFTIK LLFDENGVLM ESSN......
   15fiber  SVSLLVVKGK FSNINNNTNP NEADKQITVK LLFDANGVLK QGST......
   17fiber  NVSLIVVSGK YQYIDHATNP ..TLKSFKIK LLFDNKGVLL PSSN......
    2fiber  TVAALAV.S. ....GDLSSM TGTVASVSIF LRFDQNGVLM ENSS......
    5fiber  TVSVLAV.K. ....GSLAPI SGTVQSAHLI IRFDENGVLL NNSF......
```

FIGURE 8C

```
     4fiber   TVSVLVVRS. ..  GNLNPI TGTVSSAQVF LRFDANGV   TEHS......
  40-1fiber   TITIKGLKGA LREMNDNA.. ......LSVK LPFDNQGNLL NCA.......
    41fiber   TITIKGLKGA LREMHDNA.. ......LSLK LPFDNQGNLL NCA.......
  40-2fiber   TISIKAQKGT LL..KPTASF ......ISFV MYFYSDGTWR KNYPVFDNEG
    12fiber   IVSLVGVKGN LLNIQSTTTT ......VGVH LVFDEQGRLI TSTP.....T
     3fiber   YVTLMGASDY VNTLFKNKNV .....SINVE LYFDATGHIL PDSSSLKTDL

              551                                                600
     8fiber   ..LGKSYWNF RNQNSIMSTA YEKAIGFMPN LVAYPKPTTG SKKY...ARD
     9fiber   ..LGKSYWNF RNENSIMSTA YEKAIGFMPS LVAYPKPTAG SKKY...ARD
    15fiber   ..MDSSYWNY RSDNSNLSQP YKKAVGFMPS KTAYPKQTKP TNKEISQAKN
    17fiber   ..LDSTYWNF RSDNLTVSEA YKNAVEFMPN LVAYPKPTTG SKKY...ARD
     2fiber   ..LKKHYWNF RNGNSTNANP YTNAVGFMPN LLAYPKTQSQ T......AKN
     5fiber   ..LDPEYWNF RNGDLTEGTA YTNAVGFMPN LSAYPKSHGK T......AKS
     4fiber   ..TSKKYWGY KQGDSIDGTP YTNAVGFMPN STAYPKTQSS T......TKN
  40-1fiber   ..LESSTWRY QETNAVA... .SNALTFMPN STVYPRNKTA D......PGN
    41fiber   ..LESSTWRY QETNAVA... .SNALTFMPN STVYPRNKTA H......PGN
  40-2fiber   ILANSATWGY RQGQSANTN. VSNAVEFMPS SKRYPNEKGS E......VQN
    12fiber   ALVPQASWGY RQGQSVSTNT VTNGLGFMPN VSAYPRPNAS E......AKS
     3fiber   ELKYKQTADF .......... ..SARGFMPS TTAYPFVLPN AGTH...NEN

              601                                                650
     8fiber   IVYGNIYLGG KPHQ..PVTI KTTFNQETG. ......CEYS ITFDFSWAKT
     9fiber   IVYGNIYLGG KPDQ..PVTI KTTFNQETG. ......CEYS ITFDFSWAKT
    15fiber   KIVSNVYLGG KIDQ..PCVI IISFNEEAD. ......SDYS IVFYFKWYKT
    17fiber   IVYGNIYLGG LAYQ..PVVI KVTFNEEAD. ......SAYS ITFEFVWNKE
     2fiber   NIVSQVYLHG DKTK..PMIL TITLNGTSES TETSEVSTYS MSFTWSWESG
     5fiber   NIVSQVYLNG DKTK..PVTL TITLNGTQET GDTT.PSAYS MSFSWDWSGH
     4fiber   NIVGQVYMNG DVSK..PMLL TITLNGTDDT T.....SAYS MSFSYTWTNG
  40-1fiber   MLI....... QISP..NITF SVVYNEINS. .......GYA FTFKW.SAEP
    41fiber   MLI....... QISP..NITF SVVYNEINS. .......GYA FTFKW.SAEP
  40-2fiber   MALTYTFLQG DPNM..AISF QSIYN..HA. .....IEGYS LKFTW.RVRN
    12fiber   QMVSLTYLQG DTSK..PITM KVAFNGITS. .....LNGYS LTFMW.SGLS
     3fiber   YIFGQCYYKA SDGALFPLEV TVMLNKRLPD SRTSYVMTFL WSLNAGLAPE

              651                   672
     8fiber   .YVNVEFETT SFTFSYIAQE *.
     9fiber   .YVNVEFETT SFTFSYIAQE *.
    15fiber   .YENVQFDSS SFNFSYIAQE *.
    17fiber   .YARVEFETT SFTFSYIAQQ *.
     2fiber   KYTTETFATN SYTFSYIAQE ..
     5fiber   NYINEIFATS SYTFSYIAQE *.
     4fiber   SYIGATFGAN SYTFSYIAQQ *.
  40-1fiber   ...GKPFHPP TAVFCYITEQ *.
    41fiber   ...GKPFHPP TAVFCYITEQ *.
  40-2fiber   ...NERFDIP CCSFSYVTEQ *.
    12fiber   NYINQPFSTP SCSFSYITQE *.
     3fiber   T.TQATLITS PFTFSYIRED D*
```

FIGURE 8D

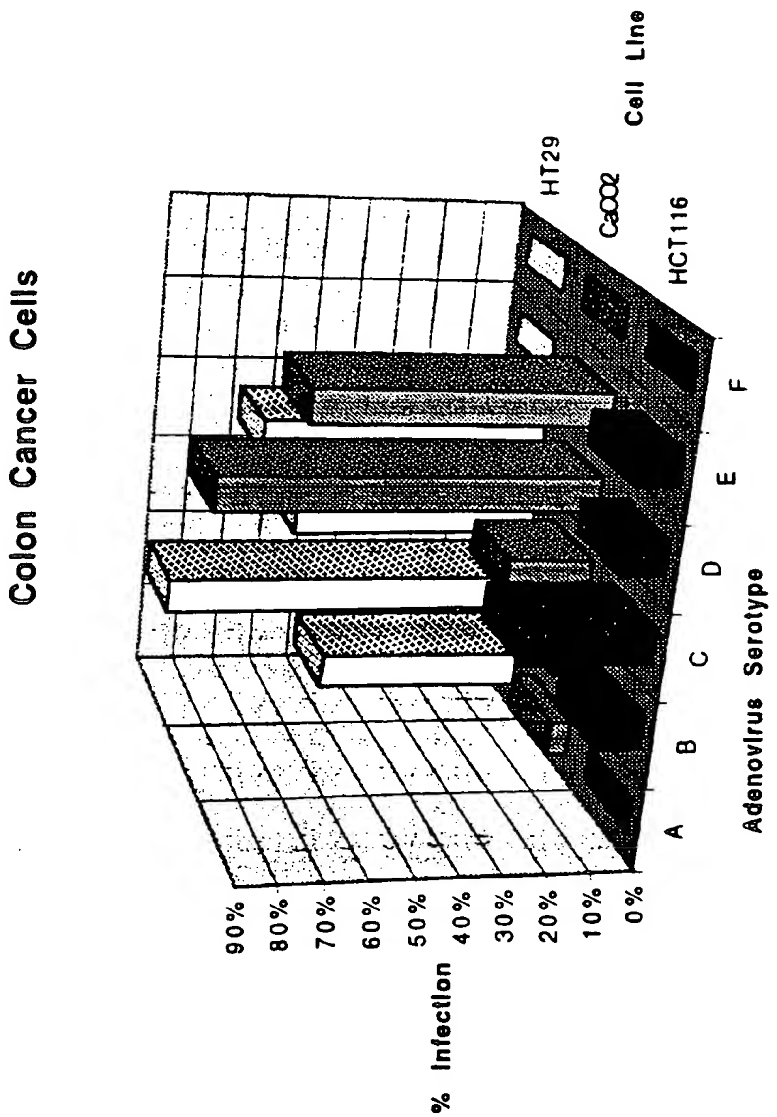
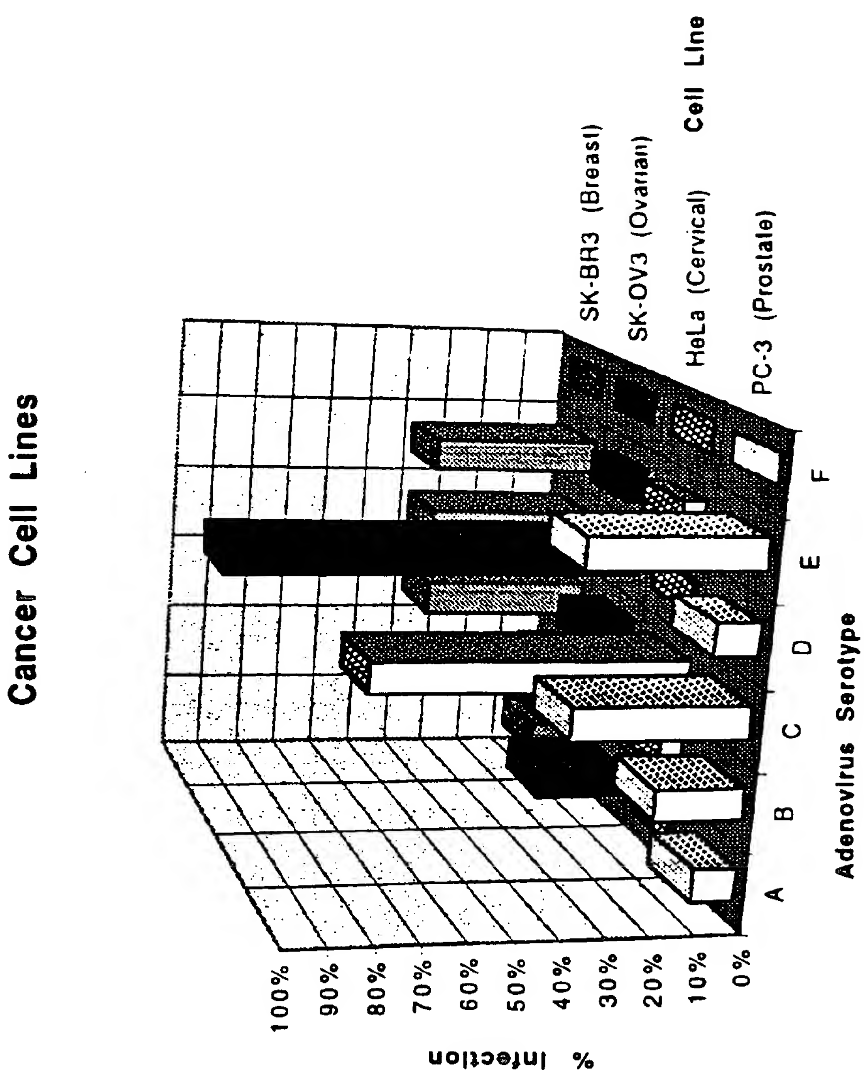

Colon Cancer Cells
% Infection
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
A
B
C
D
E
F
Adenovirus Serotype
HT29
CaCO2
HCT116
Cell Line

FIGURE 9


Cancer Cell Lines
% Infection
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
A
B
C
D
E
F
Adenovirus Serotype
SK-BR3 (Breast)
SK-OV3 (Ovarian)
HeLa (Cervical)
PC-3 (Prostate)
Cell Line

EXAMPLE 1C

INTERNATIONAL SEARCH REPORT

Internal Application No
PCT/US 97/21494

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/86 A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N A61K C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	P.W. ROELVINK ET AL.: "Comparative analysis of adenovirus fiber-cell interaction: Ad2 and Ad9 utilize the same cellular fiber receptor but use different binding strategies for attachment" JOURNAL OF VIROLOGY, vol. 70, no. 11, November 1996, AMERICAN SOCIETY FOR MICROBIOLOGY US, pages 7614-7621, XP002062100 see page 7620, last paragraph ---	1-13
A	WO 96 26281 A (GENVEC INC ;CORNELL RES FOUNDATION INC (US)) 29 August 1996 see example 7 --- -/--	1,4,6-8, 10,11

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
14 April 1998	23.04.98

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Cupido, M

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Interr nal Application No
PCT/US 97/21494

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	J. GALL ET AL: "Adenovirus type 5 and 7 capsid chimera: Fiber replacement alters receptor tropism without affecting primary immune neutralization epitopes" JOURNAL OF VIROLOGY., vol. 70, no. 4, April 1996, pages 2116-2123, XP002050655 see the whole document ---	1,4,6-8, 10,11
P,X	WO 97 12986 A (CORNELL RES FOUNDATION INC) 10 April 1997 see page 15, line 1 - line 7 -----	1,2,13

1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 97/ 21494

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: 11 to 13
because they relate to subject matter not required to be searched by this Authority, namely:
Although these claims are directed to a method of treatment of the human or animal body, the search has been carried out and based on the alleged effects of the adenoviral vector

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Interi nal Application No
PCT/US 97/21494

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9626281 A	29-08-96	AU 4980496 A	11-09-96
		CA 2213343 A	29-08-96
		EP 0811069 A	10-12-97
WO 9712986 A	10-04-97	NONE	